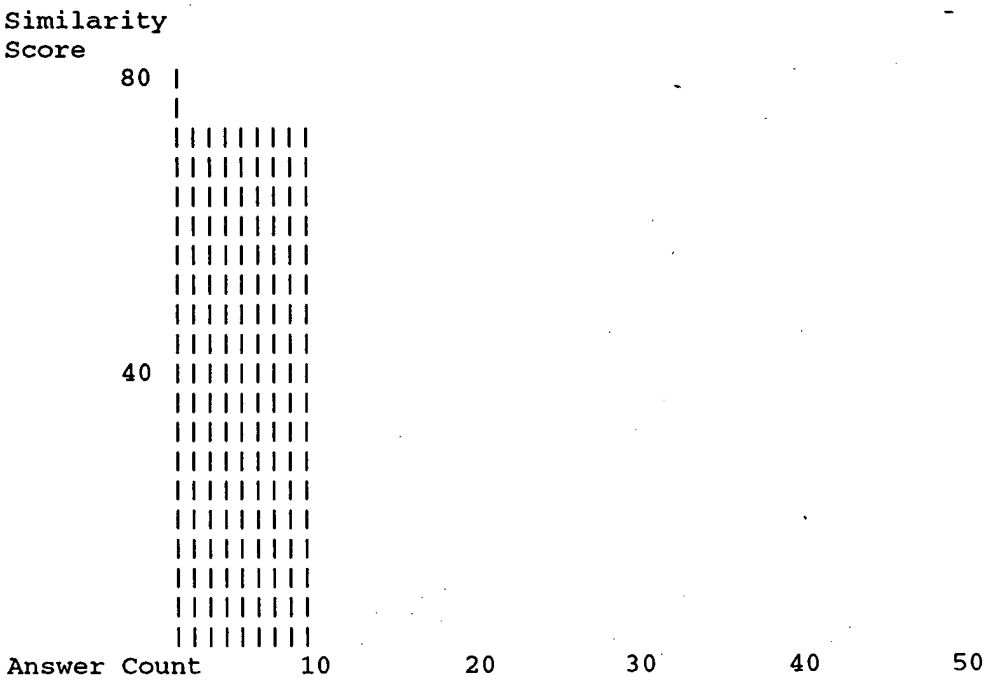
Mouse C2

MKTTTCSLLICISLLQLMVPVNTDETIEIIVENKVKELLANPANYPSTVTKTLSCTSVKTMNRWASCPAGMTATGCAC
GFACGSWEIQSGDTCNCLCLLVDWTTARCCQLS

9 ANSWERS FOUND ABOVE A THRESHOLD OF 67



L2 ANSWER 1 OF 9 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-R98208 Protein DGENE
TITLE: Cell-targetted retroviral vector particles - having envelope protein modified with targetting polypeptide
INVENTOR: Anderson W; Chiang Y L; Januszeski M; Mackrell A J; Zhao Y
PATENT ASSIGNEE: (GENE-N)GENETIC THERAPY INC
(UYSC-N) UNIV SOUTHERN CALIFORNIA
PATENT INFO: WO 9630504 A1 961003 73 pp
APPLICATION INFO: WO 96-US3908 960322
PRIORITY INFO: US 95-409648 950324
PAT. SEQ. LOC: Example 2; Page 36
DATA ENTRY DATE: 30 DEC 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-455352 [45]
DESCRIPTION: Nucleotide used in production of MSH/MoMuLV chimeric sequence
KEYWORD: Moloney murine leukaemia virus; gp70; 4070A retrovirus; retrovirus; 10A1 murine leukaemia virus; NZB-9-1 murine leukaemia virus; polytropic MX27 provirus; targetted drug delivery; gene therapy; single chain antibody; envelope protein; ss
ORGANISM: Synthetic
ABSTRACT:

Cell targetted retroviral vector particles can be used in gene therapy to deliver a heterologous gene to a target cell for expression of a heterologous polypeptide in that cell. The cell targetted retroviral vector particles comprise an envelope protein which is modified to contain a targetting polypeptide (a single chain antibody), or in the case of moloney murine leukaemia virus (MoMuLV), alpha melanotropin-stimulating hormone (MSH). Two oligonucleotides (R98207, R98208) were used to substitute sequences in MoMuLV for MSH sequences. This oligonucleotide was used to replace residues G80-P88 of MoMuLV envelope protein (See W04248)

AMINO ACID COUNTS:8 A; 0 R; 0 N; 0 D; 0 B; 17 C; 0 Q; 0 E; 0 Z; 8 G; 0 H; 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 0 S; 11 T; 0 W; 0 Y; 0 V;
SEQUENCE LENGTH: 44
SEQUENCE

```
        1 catttccgat ggtgcaagcc ggtattaacc tccctcaccc ctcg
   ALIGN Smith-Waterman score:  86
      43 aa overlap starting at 5
      scpagmtatgcacgfacgsweiqsgdtcnclcllvdwttarcc
      .: ..  .:::: :  ::.    ..  :.: :     : : ::
      tccgatggtgcaag_ccgg_tattaacctccc_____tcaccc
```

L2 ANSWER 2 OF 9 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W56120 protein DGENE
TITLE: Reducing severity of host versus graft reaction - comprises suppressing auto-immunity to heat-shock protein to prevent rejection
INVENTOR: Birk O; Cohen I R
PATENT ASSIGNEE: (YEDA)YEDA RES & DEV CO LTD
PATENT INFO: WO 9808536 A2 980305 37 pp
APPLICATION INFO: WO 97-US15294 970902
PRIORITY INFO: US 96-706209 960830
PAT. SEQ. LOC: Disclosure; Pages 28-29
DATA ENTRY DATE: 08 JUL 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 98-179175 [16]
DESCRIPTION: Protein sequence of human heat shock protein (hsp) 60
KEYWORD: Heat shock protein; hsp60; human; host versus graft reaction; HVGR; transplantation; organ; tissue; downregulatation; autoimmunity; prevention; graft rejection
ORGANISM: Homo sapiens
ABSTRACT:

The present sequence represents human heat shock protein 60 (hsp60). This protein is a stress protein expressible in all cells of the body. The severity of a host versus graft reaction (HVGR) concominant with transplantation of donor organ or tissue, can be reduced by downregulating hsp60 autoimmunity in the host. The specification also describes a method for selecting peptides for preventing or suppressing graft rejection. This method comprises treating a panel of labelled peptides with antigen-presenting cells isolated from peripheral blood lymphocytes of the candidate host, and selecting those that bind with the antigen presenting cell. hsp60, or its peptides, analogues, salts and functional derivatives can be used for downregulating hsp60 autoimmunity especially for reducing HVGR. hsp60 autoimmunity can accelerate foreign immunity and its downregulation helps prevent graft rejection

AMINO ACID COUNTS: 57 A; 22 R; 20 N; 37 D; 0 B; 3 C; 17 Q; 43 E; 0 Z; 57 G; 3 H; 43 I; 48 L; 53 K; 18 M; 10 F; 19 P; 24 S; 34 T; 1 W; 7 Y; 57 V;
SEQUENCE LENGTH: 573
SEQUENCE

```
   1 mlrlptvfrq mrpvsrvlap hltrayakdv kfgadaralm lqgvdllada
  51 vavtmgpkgr tviieqgwgs pkvtkdgvtv aksidlkdky knigaklvqd
 101 vanntneeag dgtttatvla rsiakegfek iskganpvei rrgvmlavda
 151 viaelkkqsk pvttpeeiaq vatisangdk eigniisdam kkvgrkgvit
 201 vkdgktlnde leiiegmkfd rgyispyfin tskgqkcefq dayvllsekk
 251 issiqsivpa leianahrkp lviiaedvdg ealstlvlnr lkvglqvvav
 301 kapgfgdnrk nqlkdmaiat ggavfgeegl tlnledvqph dlgkvgeviv
 351 tkddamllkg kgdkaqiekr iqeiieqldv ttseyekekl nerlaklsdg
 401 vavlkvggts dvevnekkdr vtdalnatra aveegivlgg gcallrcipa
 451 ldsltpaned qkigieiikr tlkipamtia knagvegsli vekimqssse
 501 vgydamagdf vnmvekgiid ptkvvrtall daagvasllt taevvvteip
 551 keekdpgmga mggmgggmgg gmf
```

ALIGN Smith-Waterman score: 81
38 aa overlap starting at 442

```
csllicisllqlmvpvntdetieiivenkvkellanpa
:.:: ::  :. ..:.: :.  .: .:  .:. :  ::
callrcipaldsltpanedq__kigiei_ikrtlkipa
```

SEARCHED ON 26 OCT 1998
FILE LAST UPDATED: 18 OCT 1998 <19981018/UP>

L2 ANSWER 3 OF 9 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W12345 protein DGENE
TITLE: New peptide(s) derived from human heat-shock protein 60 - used for early diagnosis, prevention and treatment of insulin-dependent diabetes mellitus
INVENTOR: Abulafia R; Bockova J; Cohen I R; Elias D
PATENT ASSIGNEE: (YEDA)YEDA RES & DEV CO LTD
PATENT INFO: WO 9701959 A1 970123 49 pp
APPLICATION INFO: WO 96-US11375 960701
PRIORITY INFO: IL 95-114407 950630
PAT. SEQ. LOC: Disclosure; Page 28-29
DATA ENTRY DATE: 13 NOV 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-108693 [10]
DESCRIPTION: Human heat-shock protein 60
KEYWORD: Heat-shock protein; hsp; hsp60; insulin-dependent diabetes mellitus; IDDM
ORGANISM: Homo sapiens
ABSTRACT:

The peptides given in W12346 to W12358 are derived from human hsp60 (W12345) and are useful for early diagnosis of IDDM by detecting, in the blood or urine, antibodies or T-cells immunologically reactive with human hsp60 (presence of which indicates high probability of diabetes or its subsequent development). Other peptides (W12359 to W12361) were shown not to be as effective

AMINO ACID COUNTS: 57 A; 22 R; 20 N; 37 D; 0 B; 3 C; 17 Q; 43 E; 0 Z; 57 G; 3 H; 43 I; 48 L; 53 K; 18 M; 10 F; 19 P; 24 S; 34 T; 1 W; 7 Y; 57 V;
SEQUENCE LENGTH: 573
SEQUENCE

```
     1 mlrlptvfrq mrpvsrvlap hltrayakdv kfgadaralm lqgvdllada
    51 vavtmgpkgr tviieqgwgs pkvtkdgvtv aksidlkdky knigaklvqd
   101 vanntneeag dgtttatvla rsiakegfek iskganpvei rrgvmlavda
   151 viaelkkqsk pvttpeeiaq vatisangdk eigniisdam kkvgrkgvit
   201 vkdgktlnde leiiegmkfd rgyispyfin tskgqkcefq dayvllsekk
   251 issiqsivpa leianahrkp lviiaedvdg ealstlvlnr lkvglqvvav
   301 kapgfgdnrk nqlkdmaiat ggavfgeegl tlnledvqph dlgkvgeviv
   351 tkddamllkg kgdkaqiekr iqeiieqldv ttseyekekl nerlaklsdg
   401 vavlkvggts dvevnekkdr vtdalnatra aveegivlgg gcallrcipa
   451 ldsltpaned qkigieiikr tlkipamtia knagvegsli vekimqssse
   501 vgydamagdf vnmvekgiid ptkvvrtall daagvasllt taevvvteip
   551 keekdpgmga mggmgggmgg gmf
```

ALIGN Smith-Waterman score: 81
38 aa overlap starting at 442

```
csllicisllqlmvpvntdetieiivenkvkellanpa
:.:: ::  :. ..:.: :.  .: .:  .:. :  ::
callrcipaldsltpanedq__kigiei_ikrtlkipa
```

L2 ANSWER 4 OF 9 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W01657 Protein DGENE
TITLE: Compsns. for treating or preventing insulin-dependent diabetes - based on T cells specific for 65 kD heat-shock protein
INVENTOR: Cohen I R; Elias D; Markovits D
PATENT ASSIGNEE: (YEDA)YEDA RES & DEV CO LTD
PATENT INFO: US 5578303 A 961126 29 pp
APPLICATION INFO: US 89-322864 890314
PRIORITY INFO: US 91-751448 910829
US 89-322864 890314
US 89-371249 890626
US 90-493127 900314
US 93-151052 931112
PAT. SEQ. LOC: Disclosure; Fig 3A-3B
DATA ENTRY DATE: 25 APR 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-020369 [02]
CROSS REFERENCES: N-PSDB: 97N-T58403
DESCRIPTION: Human heat shock protein 65
KEYWORD: Heat shock protein 65; Hsp65; insulin-dependent diabetes mellitus; IDDM; autoimmune disease; diagnosis; therapy; T cell; vaccine
ORGANISM: Homo sapiens
ABSTRACT:

The human heat shock protein 65 (Hsp65) (W01657) is expressed in the islets of the pancreas. The T cell response to Hsp65 is associated with the development of insulin-dependent diabetes mellitus (IDDM). A method for detecting the existence of, a tendency to develop, or the initiation of a process leading to IDDM involves detecting the presence of Hsp65 or antibodies or T cells reactive with the protein. Hsp65, when administered to a tolerogenic carrier, can be used to prevent or treat IDDM prior to development of clinical symptoms. Attenuated T cells can be used to vaccinate against autoimmunity to Hsp65 and to abort IDDM

AMINO ACID COUNTS: 56 A; 22 R; 21 N; 35 D; 0 B; 3 C; 18 Q; 40 E; 0 Z; 58 G; 2 H; 41 I; 49 L; 54 K; 18 M; 11 F; 19 P; 24 S; 34 T; 1 W; 7 Y; 60 V;
SEQUENCE LENGTH: 573
SEQUENCE

```
      1 mlrlptvfrq mrpvsrvlap hltrayakdv kfgadaralm lqgvdllada
     51 vavtmgpkgr tviieqswgs pkvtkdgvtv aksidlkdky knigaklvqd
    101 vanntnegag dgtttatvla rsiakegfek iskganpvei rrgvmlavda
    151 viaetkkqsk pvttpeeiaq vatisangdk eigniisdam kkvgrkgvit
    201 vkdgktlnde leiiegmkfd rgyispyfin tskgqkcefq dayvllsekk
    251 issiqsivpa leianlvlnr lkvglqvvav kapgflvlnr lkvglqvvav
    301 kapgfgdnrk nqlkdmaiat ggavfgeegl tlnledvqph dlgkvgeviv
    351 tkddamllkg kgdkaqiekr iqeiieqldv ttseyekekl nerlaklsdg
    401 vavlkvggts dvevnekkdr vtdalnatra aveegivlgg gcallrcipa
    451 ldsltpaned qkigieiikr tlkipamtia knagvegsli vekimqssse
    501 vgydamagdf vnmvekgiid ptkvvrtall daagvasllt taevvvteip
    551 keekdpgmga mggmgggmgg gmf
```

FEATURE TABLE:

Key	Location	Qualifier	
Peptide	1..26	label	Leader_peptide
		note	"putative mitochondrial targeting sequence"
Protein	27..573	label	Mat_protein
		note	"amino acid residues 266-285 differ from the translated sequence (LVLNRLKVGLQVVAVKAPGF)
Region	557..572	note	"keratin-like region contg. Gly-Gly-Met repeats""

ALIGN Smith-Waterman score: 81
38 aa overlap starting at 442

```
cslliclsllqlmvpvntdetieiivenkvkellanpa
:.:: ::  :. ..:.: :.  .: .:  .:. :  ::
callrcipaldsltpanedq__kigiei_ikrtlkipa
```

L2 ANSWER 5 OF 9 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R67385 Protein DGENE
TITLE: DNA from Helicobacter pylori and Helicobacter felis - used to develop prods. for detection, treatment and prevention of Helicobacter infection
INVENTOR: Ferrero R; Labigne A; Suerbaum S; Thiberge J
PATENT ASSIGNEE: (INRM) INST NAT SANTE & RECH MEDICALE
(INSP) INST PASTEUR
PATENT INFO: WO 9426901 A 941124 168 pp
APPLICATION INFO: WO 94-EP1625 940519
PRIORITY INFO: EP 93-401309 930519
WO 93-EP3259 931119
PAT. SEQ. LOC: Disclosure; Fig. 7A(i-vii)
DATA ENTRY DATE: 22 JUN 1995 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 95-006797 [01]
DESCRIPTION: Mitochondrial protein P1
KEYWORD: Urease; immunogen; vaccine; diagnostic; heat shock protein; HSP; GroEL-like protein; Helicobacter felis
ORGANISM: Homo sapiens
ABSTRACT:

The sequence of the Helicobacter pylori heat shock protein A (given in R67374) was compared to that of other GroEL-like proteins from Legionella pneumophila (R67381), Escherichia coli (R67382), Chlamydia psittaci (R67383), Mycobacterium leprae (R67384) and human mitochondrial protein P1 (R67385), and regions of homology were identified

AMINO ACID COUNTS: 52 A; 17 R; 21 N; 37 D; 0 B; 3 C; 17 Q; 40 E; 0 Z; 56 G; 1 H; 41 I; 46 L; 52 K; 16 M; 10 F; 16 P; 26 S; 31 T; 1 W; 7 Y; 57 V;
SEQUENCE LENGTH: 547
SEQUENCE

```
   1 ymadvkfgad aralmlqgvd lladavavtm gpkgrtviie qswgspkvtk
  51 dgvtvaksid lkdkykniga klvqdvannt neeagdgttt atvlarsiak
 101 egfekiskga npveirrgvd lavdaviael kkqskpvttp eeiaqvatis
 151 angdkeigni isdamkkvgr kgvitvkdgk tlndeleiie gmkfdrgyis
 201 pyfintskgq kcefqdayvl lsekkissiq sivpaleian lvlnrlkvgl
 251 qvvavkapgf lvlnrlkvgl qvvavkapgf gdnrknqlkd maistggsvf
 301 geegltlnle dvqphdlgkv gevivtkdda mllkgkgdka qiekriqeii
 351 eqldvttsey ekeklnerla klsdgvavlk vggtsdvevn ekkdrvtdal
 401 natraaveeg ivlgggcall rcipaldslt panedqkigi eiikrtlkip
 451 amtiaknagv dgslivekim qsssevgyda magdfvnmve kgiidptkvv
 501 rtalldaasv aslttaevv vteipeekdp gmgamggmgg gmgggmf
```

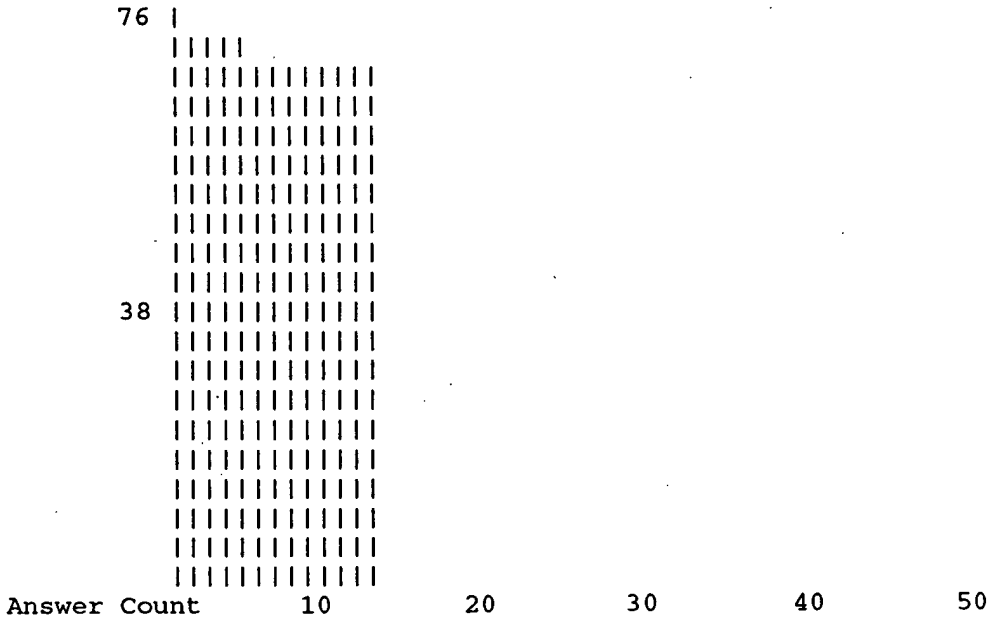
ALIGN Smith-Waterman score: 81
38 aa overlap starting at 417

```
csllicisllqlmvpvntdetieiivenkvkellanpa
:.:: ::  :. ..:.: :.  .: .:  .:. :  ::
callrcipaldsltpanedq__kigiei_ikrtlkipa
```

Mouse C2b

MKTTTCSLLICISLLQLMVPVNTEGTLESIVEKKVKELLANRDDCPSTVTKTFSCTSITASGRLASCPSGMTVTGCAC
GYGCGSWDIRDGNTCHCQCSTMDWATARCCQLA

13 ANSWERS FOUND ABOVE A THRESHOLD OF 67



L4 ANSWER 1 OF 13 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-W05534 Peptide DGENE
TITLE: New gC1q receptor-based, HIV-1 gp 120 binding peptide(s) - for preventing and treating HIV-1 infection
INVENTOR: Fung M S C; Kim Y W; Sun B N V; Sun C R Y; Yu L
PATENT ASSIGNEE: (TANO-N)TANOX BIOSYSTEMS INC
PATENT INFO: WO 9630400 A1 961003 53 pp
APPLICATION INFO: WO 96-US3905 960322
PRIORITY INFO: US 95-410360 950324
PAT. SEQ. LOC: Disclosure; Page 46-48
DATA ENTRY DATE: 17 JAN 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-455274 [45]
CROSS REFERENCES: N-PSDB: 96N-T41465
DESCRIPTION: gC1q receptor
KEYWORD: gC1q receptor; gC1q-R; human immunodeficiency virus type 1; HIV-1; gp120; immunogen; vaccine; therapy; diagnosis
ORGANISM: Homo sapiens
ABSTRACT:

The gC1q receptor (gC1q-R) (W05534), a receptor for C1q complement, binds to HIV-1 gp120 and neutralises the infectivity of HIV-1. The binding site for gp120 has been identified (see also W05532). The receptor exists on a variety of cell types, including B cells, T cells, monocytes, macrophages, neutrophils, eosinophils, platelets, fibroblasts and endothelial, liver, neural and smooth muscle cells. Recombinant, mature gC1q-R can be produced in transformed host cells (see also T41465). It is useful for detecting or quantifying HIV-1 gp120, HIV-1 virions or HIV-1 infected cells, and can also be used to treat or prevent HIV-1 infection and to raise antibodies of diagnostic or therapeutic value

AMINO ACID COUNTS:17 A; 15 R; 9 N; 20 D; 0 B; 7 C; 10 Q; 29 E; 0 Z;
20 G; 5 H; 9 I; 31 L; 16 K; 3 M; 13 F; 17 P; 20 S;
16 T; 3 W; 4 Y; 18 V;
SEQUENCE LENGTH: 282
SEQUENCE

```
      1 mlpllrcvpr vlgssvaglr aaapaspfrq llqpaprlct rpfgllsvra
     51 gserrpgllr prgpcacgcg cgslhtdgdk afvdflsdei keerkiqkhk
    101 tlpkmsggwe lelngteakl vrkvagekit vtfninnsip ptfdgeeeps
    151 qgqkveeqep eltstpnfvv eviknddgkk alvldchype devgqedeae
    201 sdifsirevs fqstgesewk dtnytlntds ldwalydhlm dfladrgvdn
    251 tfadelvels talehqeyit fledlksfvk sq
```

FEATURE TABLE:

Key	Location	Qualifier	
Peptide	1..73	label	Pre-propeptide
Protein	74..282	label	Mat_protein
Modified_site	114	label	Glycosylation
		note	"potential N-glycosylation site"
Modified_site	136	label	Glycosylation
		note	"potential N-glycosylation site"

SEARCHED ON 26 OCT 1998
FILE LAST UPDATED: 18 OCT 1998 <19981018/UP>

```
Modified_site  |223      |label    |Glycosylation
               |         |note     |"potential N-glycosylation
               |         |         |site"
Binding_site   |239..250 |label    |gp120_binding_site
               |         |note     |"Claim 1"

ALIGN Smith-Waterman score:  85
      47 aa overlap starting at 36
      pstvtktfsctsitasgrlascpsgmtvtg_cacgygcgswdirdgn
      :   :. :.  :. : :     :. .   : :::: ::::    ::.
      prlctrpfgllsvra_gse_rrpgllrprgpcacgcgcgslh_tdgd
```

L4 ANSWER 2 OF 13 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-R91446 Protein DGENE
TITLE: Cloning of cDNA encoding cell surface antigen - useful for isolation of diagnostic and therapeutic proteins
INVENTOR: Aruffo A; Seed B
PATENT ASSIGNEE: (GEHO)GEN HOSPITAL CORP
PATENT INFO: US 5506126 A 960409 79 pp
APPLICATION INFO: US 88-160416 880225
PRIORITY INFO: US 92-983647 921201
US 88-160416 880225
US 89-379076 890713
US 90-553759 900713
US 93-139273 931018
PAT. SEQ. LOC: Example 16; Column 85-88
DATA ENTRY DATE: 31 OCT 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-200279 [20]
CROSS REFERENCES: N-PSDB: 96N-T14726
DESCRIPTION: Human CD53 antigen
KEYWORD: Cell surface antigen; cloning; immunoselection; immunotherapy; therapy; diagnosis; vector; COS; CD53; lymphocyte
ORGANISM: Homo sapiens
ABSTRACT:

Human antigen CD53 (R91446) is a type III integral membrane protein that may be involved in the transport of factors essential for cell proliferation. Its amino acid sequence was deduced from a cDNA clone (T14726) obtd. using a novel immunoselection cloning technique. CD53 was expressed in transfected COS cells. Anti-CD53 antibodies are a useful tool for the identification of haematopoietic neoplasms, and may prove helpful for identifying morphologically poorly defined cells

AMINO ACID COUNTS: 10 A; 3 R; 11 N; 6 D; 0 B; 12 C; 5 Q; 6 E; 0 Z; 17 G; 5 H; 22 I; 33 L; 10 K; 6 M; 16 F; 4 P; 18 S; 10 T; 4 W; 7 Y; 14 V;
SEQUENCE LENGTH: 219
SEQUENCE

```
      1 mgmsslkllk yvlfffnllf wicgccilgf giyllihnnf gvlfhnlpsl
     51 tlgnvfvivg siimvvaflg cmgsikenkc llmsffilll iillaevtla
    101 illfvyeqkl neyvakgltd sihryhsdns tkaawdsiqs flqccgingt
    151 sdwtsgppas cpsdrkvegc yakarlwfhs nflyigiiti cvcvievlgm
    201 sfaltlncqi dktsqtigl
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	8..36	label	Hydrophobic_domain
Domain	55..75	label	Hydrophobic_domain
Domain	81..106	label	Hydrophobic_domain
Modified_site	149..151	label	Glycosylation_site
Modified_site	228..230	label	Glycosylation_site
Domain	182..206	label	Hydrophobic_domain

```
ALIGN Smith-Waterman score:  73
      96 aa overlap starting at 75
      mktttcsll_icisllqlmvpvntegtlesivekkvkella_____________nrddcp
      .: . : :. . : :: ...   : . :  . :.:..: .:             :
      ikenkcllmsffillliillaevtlaillfvyeqklneyvakgltdsihryhsdnstkaa
      _stvtktfsctsi___t__asgrlascpsgmtvtgc
       ... . ..: .:   .  .::  :::::   : ::
      wdsiqsflqccgingtsdwtsgppascpsdrkvegc
```

L4 ANSWER 3 OF 13 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 94P-R63790 Protein DGENE
TITLE: New xylanase enzymes from Aspergillus aculeatus - used for degrading plant cell wall components, e.g. in the prepn. of feed, in baking and in prepn. of pulp or paper
INVENTOR: Andersen L N; Christgau S; Dalboge H; Heldt-hansen H P; Jacobsents; Kauppinen M S; Kofod L V; Mullertz A; Munk N; Si J Q
PATENT ASSIGNEE: (NOVO)NOVO-NORDISK AS
PATENT INFO: WO 9421785 A 940929 80 pp
APPLICATION INFO: WO 94-DK88 940302
PRIORITY INFO: DK 93-268 930310
DK 93-1151 931014
PAT. SEQ. LOC: Claim 10; Page 57
DATA ENTRY DATE: 07 JUN 1995 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 94-317006 [39]
CROSS REFERENCES: N-PSDB: 94N-Q74637
DESCRIPTION: Aspergillus aculeatus xylanase II
KEYWORD: Xylanase II; Aspergillus aculeatus; alpha-arbino-pyranosidase; brewing; paper pulp; food preparation; plant cell wall degradation
ORGANISM: Aspergillus aculeatus
ABSTRACT:
Q74637 encodes R63790 Aspergillus aculeatus xylanase II, which degrades plant cell wall components and reduces the viscosity of plant cell wall derived material. These properties are useful in the production of dough and baked products; in the preparation of feed, food, beer, wine, pulp and paper; and for the separation of cereal components. In addition xylanase II exhibits alpha-arbino-pyranosidase activity, and it can also be used in the production of antibodies
AMINO ACID COUNTS: 54 A; 6 R; 22 N; 21 D; 0 B; 8 C; 20 Q; 10 E; 0 Z; 39 G; 9 H; 17 I; 38 L; 18 K; 4 M; 13 F; 16 P; 28 S; 64 T; 11 W; 20 Y; 33 V; 2 Others;
SEQUENCE LENGTH: 453
SEQUENCE

```
   1 mvgllsitaa laatvlpniv savgldqaav akglqyfgta tdnpeltdip
  51 yvtqlnntad fgqitpgnsm kwdatepsqg tftftkgdvi adlaegngqy
 101 lrchtlvwyn qlpswvtsgt wtnatltaal knhitnvvsh ykgkclhwdv
 151 vnealnddgt yrtnifytti geayipiafa aaaaadpdak lfyndynley
 201 ggakaasara ivqlvknaga kidgvglqah fsvgtvpsts slvsvlqsft
 251 algvevayte advrillptt attlaqqssd fqalvqscvq ttgcvgftiw
 301 dwtdkyswvp stfsgygaal pwdenlvkkp ayngllagmg vtvttttttt
 351 tatatgkttt tttgatstgt taahwgqcgg lnwsgptaca tgytctyvnd
 401 yysqclxsia qpkpagvlai qsvrfiyhnt qnslldlxkk ktlehtggrs
 451 smh
```

FEATURE TABLE:

Key	Location	Qualifier	
Misc_difference	442	note	"corresponding codon codes Lys"
Misc_difference	443..453	note	"no corresponding codons"

```
ALIGN Smith-Waterman score:  89
   64 aa overlap starting at 342
   tvtktfsctsitasgrlascpsgmtvtg_cacgyg_cgswdirdgntchc__qcstmdwa
   ::: : . :. ::.:. ..  .: : ::  :  .: ::. .    ..:     :. ..
   tvttttttttatatgkttttttgatstgttaahwgqcgglnwsgptacatgytctyvndy
   tarc
    ..:
   ysqc
```

L4 ANSWER 4 OF 13 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 93P-R45352 protein DGENE
TITLE: New polypeptide(s) from venom of Plectreurys tristis - used as insecticidal agents which are paralytic and/or toxic to insects at low concns
INVENTOR: Leisy D J; Quistad G B; Skinner W S
PATENT ASSIGNEE: (SANO)SANDOZ AG
(SANO) SANDOZ PATENT GMBH
(SANO) SANDOZ-ERFINDUNGEN VERW.GES MBH
PATENT INFO: EP 556160 A 930818 50 pp
APPLICATION INFO: EP 93-810078 930208
PRIORITY INFO: US 92-837194 920211
PAT. SEQ. LOC: Claim 1; Page 42
DATA ENTRY DATE: 04 FEB 1994 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 93-260398 [33]
DESCRIPTION: Arachnoidal plectoxin polypeptide G
KEYWORD: Primitive Hunting Spider; insecticidal; plectoxin; arachnid
ORGANISM: Plectreurys tristis
ABSTRACT:
This polypeptide corresponds to a plectoxin sequence isolated from the Primitive Hunting Spider (Plectreurys tristis). The plectoxins have insecticidal properties and are used against Lepidopteran species. See also R38840 and R45351
AMINO ACID COUNTS:3 A; 2 R; 2 N; 1 D; 0 B; 10 C; 1 Q; 1 E; 0 Z; 7 G; 0 H; 1 I; 3 L; 7 K; 0 M; 2 F; 0 P; 5 S; 1 T; 1 W; 0 Y; 2 V;
SEQUENCE LENGTH: 49
SEQUENCE
1 gckgflvkcd snseccktai vkgkkkqlsc lcgawgagcs csfrcgnrc
ALIGN Smith-Waterman score: 76
32 aa overlap starting at 16
c_tsitasg_rlascpsgmtvtgcacgygcgs
: :.:. . . :: : .::.:.. ::.
cktaivkgkkkqlsclcgawgagcscsfrcgn

L4 ANSWER 5 OF 13 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 92P-R20818 Protein DGENE
TITLE: New CD53 cell surface antigen and DNA encoding it - for immuno-therapy and diagnosis of haematopoietic neoplasms, etc
INVENTOR: Seed B; Aruffo A; Amiot M
PATENT ASSIGNEE: (GEHO-N)GEN HOSPITAL CORP
PATENT INFO: WO 9201049 A 920123 160 pp
APPLICATION INFO: WO 90-US4986 900715
PRIORITY INFO: US 90-553759 900713
PAT. SEQ. LOC: Claim 4; Page 123
DATA ENTRY DATE: 21 MAY 1992 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 92-056864 [07]
CROSS REFERENCES: N-PSDB: 92N-Q21187
DESCRIPTION: CD53 haematopoietic antigen
KEYWORD: Rapid immunoselection cloning technique; cell surface antigen; haematopoietic neoplasm; type III integral membrane protein
ORGANISM: Homo sapiens
ABSTRACT:

A cDNA clone encoding CD53 was obtained using the rapid immunoselection cloning method (see e.g Q21164 for description of method). The cDNA libraries were prepared from the promyelocytic tumour cell line HL60 and from peripheral blood lymphocytes and transfected into COS cells. The first of the four predicted hydrophobic regions is atypically long for either a signal sequence or a simple transmembrane alpha-helix. Both cysteine and glycine have been found to precede the signal cleavage site (Von Heijne, Nucleic Acid Res. 14:4683 (1986)) and the presence of 3 cysteines and a glycine in the middle of the first hydrophobic region suggests that the N-terminus of the mature CD53 protein begins there

AMINO ACID COUNTS:10 A; 3 R; 11 N; 6 D; 0 B; 12 C; 5 Q; 6 E; 0 Z;
17 G; 5 H; 22 I; 33 L; 10 K; 6 M; 16 F; 4 P; 18 S;
10 T; 4 W; 7 Y; 14 V;
SEQUENCE LENGTH: 219
SEQUENCE

```
  1 mgmsslkllk yvlfffnllf wicgccilgf giyllihnnf gvlfhnlpsl
 51 tlgnvfvivg siimvvaflg cmgsikenkc llmsffilll iillaevtla
101 illfvyeqkl neyvakgltd sihryhsdns tkaawdsiqs flqccgingt
151 sdwtsgppas cpsdrkvegc yakarlwfhs nflyigiiti cvcvievlgm
201 sfaltlncqi dktsqtigl
```

FEATURE TABLE:

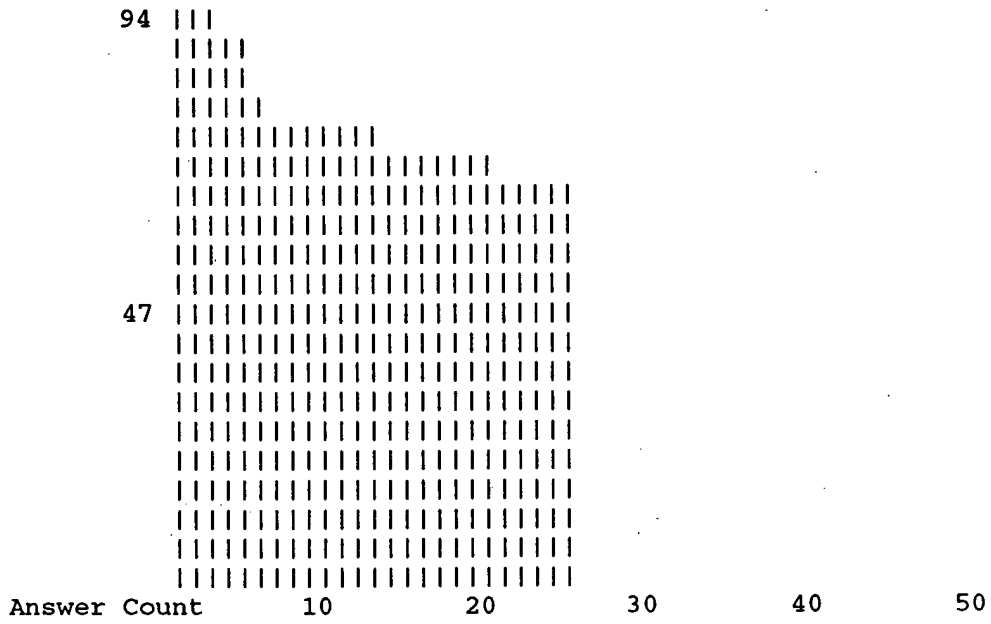
Key	Location	Qualifier	
Region	8..36	label	hydrophobic
Region	55..75	label	hydrophobic
Region	81..106	label	hydrophobic
Region	182..206	label	hydrophobic
Modified_site	149..151	label	N-linked_glycosylation
		note	"putative"
Modified_site	168..170	label	N-linked_glycosylation
		note	"putative"

```
ALIGN Smith-Waterman score:  73
      96 aa overlap starting at 75
      mktttcsll_icisllqlmvpvntegtlesivekkvkella_____________nrddcp
      .: . : :. . : :: ...   : . :  . :.:..: .:              :
      ikenkcllmsffillliillaevtlaillfvyeqklneyvakgltdsihryhsdnstkaa
      _stvtktfsctsi___t__asgrlascpsgmtvtgc
       ... . ..: .:  . .::  :::::   : ::
      wdsiqsflqccgingtsdwtsgppascpsdrkvegc
```

Human C10

MGPSSCLLLILIPLLQLINPGSTQCSLDSVMDKKIKDVLNSLEYSPSPISKKLSCASVKSQGRPSSCPAGMAVTGCAC
GYGCGSWDVQLETTCHCQCSVVDWTTARCCHLT

25 ANSWERS FOUND ABOVE A THRESHOLD OF 67



L6 ANSWER 1 OF 25 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W60854 Protein DGENE
TITLE: New bovine polypeptide that activates mammalian B cell(s) - used e.g. to treat T cell immunodeficiency or allergy, as vaccine adjuvant, as T cell surrogate for infants, and for monoclonal antibody production, also specific antibodies for treating B cell hyperactivity
INVENTOR: Alizadeh-Khiavi K; Filipp D; Julius M H
PATENT ASSIGNEE: (WELL-N)WELLESLEY HOSPITAL FOUND
PATENT INFO: WO 9822580 A2 980528 64 pp
APPLICATION INFO: WO 97-CA880 971118
PRIORITY INFO: US 96-746883 961118
PAT. SEQ. LOC: Claim 14; Fig 7
DATA ENTRY DATE: 01 OCT 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 98-312466 [27]
CROSS REFERENCES: N-PSDB: 98N-V37228
DESCRIPTION: Human CD14 protein
KEYWORD: CD14; B cell activator; bovine lactation-associated immunotropic protein; LAIT; T cell immunodeficiency; X-linked hyper-IgM syndrome; allergy; common variable immunodeficiency; X-linked agammaglobulinaemia; vaccine; infant feeding formulae; human
ORGANISM: Homo sapiens
ABSTRACT:

This sequence is the human CD14 protein of the invention. The CD14 protein was used to isolate the bovine CD14 of the invention, which is able to activate mammalian B cells. The protein is also known as bovine lactation-associated immunotropic protein (LAIT), and is used to activate B cells, particularly in humans. Particularly it is administered to subjects: (a) with CD40 negative or deficient B cells; (b) suffering from T cell immunodeficiency (e.g. X-linked hyper-IgM syndrome, common variable immunodeficiency or X-linked agammaglobulinaemia) or allergy (i.e. with CD40 ligand negative or defective T cells); or (c) to induce growth and differentiation of B cells to highly productive Ig secreting cells. Particular applications are in infant feeding formulae (as immunostimulant) and as adjuvant in vaccines (optionally with bovine CD14 coupled to the antigen). The DNA sequences are also used to enrich mammalian B cells secreting a monoclonal antibody (MAb) of particular antigenic specificity, by activating cells with sub-optimal amount of the DNA in combination with the antigen. The enriched B cells are then used to produce hybridomas that produce specific MAb. Antibodies raised against human CD14 are used to reduce/inhibit activity of B cells that are hyperactivated by high serum levels of CD14. Bovine CD14 stimulates growth (induce DNA synthesis) in resting murine spleen cells and is 200 times more effective than lipopolysaccharide (LPS), with the effect unaffected by presence of serum. It also induces Ig secretion and a partial isotype switch from IgM to IgG, in absence of T cells

AMINO ACID COUNTS: 43 A; 22 R; 15 N; 16 D; 0 B; 10 C; 14 Q; 20 E; 0 Z; 24 G; 7 H; 4 I; 62 L; 9 K; 6 M; 10 F; 30 P; 29 S; 18 T; 5 W; 3 Y; 28 V;
SEQUENCE LENGTH: 375

SEQUENCE

```
        1 merascllll llplvhvsat tpepceldde dfrcvcnfse pqpdwseafq
       51 cvsaveveih agglnlepfl krvdadadpr qyadtvkalr vrrltvgaaq
      101 vpaqllvgal rvlaysrlke ltledlkitg tmpplpleat glalsslrlr
      151 nvswatgrsw laelqqwlkp glkvlsiaqa hspafsyeqv rafpaltsld
      201 lsdnpglger glmaalcphk fpaiqnlalr ntgmetptgv caalaaagvq
      251 phsldlshns lratvnpsap rcmwssalns lnlsfagleq vpkglpaklr
      301 vldlscnrln rapqpdelpe vdnltldgnp flvpgtalph egsmnsgvvp
      351 acarstlsvg vsgtlvllqg argfa
```

FEATURE TABLE:

Key	Location	Qualifier	
Misc_difference	87	note	"encoded by TGC"

ALIGN Smith-Waterman score: 94
58 aa overlap starting at 1

```
mgpsscllliliplqlinpgstqcsldsvmdkkikdvlnsleyspspiskklscasv
:  .:::::.:.::...       : ::   :. .. : :  : .:.  :. ..:.:.
merascllllllplvhvsattpepceld___dedfrcvcnfsepqpd_wseafqcvsa
```

L6 ANSWER 2 OF 25 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W41693 Protein DGENE
TITLE: Assay for lipo:poly:saccharide binding inhibitors - useful in the treatment of septic shock and other lipo:poly:saccharide-mediated disorder(s)
INVENTOR: Mintz D N; Tobias P; Ulevitch R
PATENT ASSIGNEE: (SCRI)SCRIPPS RES INST
PATENT INFO: US 5705398 A 980106 21 pp
APPLICATION INFO: US 94-205719 940302
PRIORITY INFO: US 94-205719 940302
PAT. SEQ. LOC: Claim 9; Columns 19-22
DATA ENTRY DATE: 01 MAY 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 98-086145 [08]
CROSS REFERENCES: N-PSDB: 98N-V05505
DESCRIPTION: Human CD14
KEYWORD: Identification; binding inhibitor; lipopolysaccharide; LPS; CD14; binding protein; LBP; monocyte receptor; treatment; septic shock; human
ORGANISM: Homo sapiens
ABSTRACT:

The present sequence was used in the development of a novel method for identifying a compound that inhibits the binding of a lipopolysaccharide (LPS) to a LPS-binding protein (LBP), or LBP-dependent binding of LPS to monocyte receptor CD14. The method comprises measuring the fluorescence emitted by a reaction mixture containing fluoresceinated LPS, isolated LBP and optionally CD14 in the presence and absence of the compound, and identifying the compound as an inhibitor if the fluorescence emitted by the reaction mixture containing the compound is less than that emitted by the reaction mixture that does not contain the compound. The method can be used to identify drugs useful for treating septic shock and related LPS-mediated disorders

AMINO ACID COUNTS: 42 A; 23 R; 15 N; 16 D; 0 B; 11 C; 14 Q; 20 E; 0 Z; 24 G; 7 H; 4 I; 62 L; 9 K; 6 M; 10 F; 30 P; 29 S; 18 T; 5 W; 2 Y; 28 V;
SEQUENCE LENGTH: 375
SEQUENCE

```
      1 merascllll llplvhvsat tpepceldde dfrcvcnfse pqpdwseafq
     51 cvsaveveih agglnlepfl krvdadrdpr qyadtvkalr vrrltvgaaq
    101 vpaqllvgal rvlaysrlke ltledlkitg tmpplpleat glalsslrlr
    151 nvswatgrsw laelqqwlkp glkvlsiaqa hspafsceqv rafpaltsld
    201 lsdnpglger glmaalcphk fpaiqnlalr ntgmetptgv caalaaagvq
    251 phsldlshns lratvnpsap rcmwssalns lnlsfagleq vpkglpaklr
    301 vldlscnrln rapqpdelpe vdnltldgnp flvpgtalph egsmnsgvvp
    351 acarstlsvg vsgtlvllqg argfa
```

ALIGN Smith-Waterman score: 94
58 aa overlap starting at 1

```
mgpsscllliliplllqlinpgstqcsldsvmdkkikdvlnsleyspspiskklscasv
:  .:::::.:.::...       : ::   :. .. : :  : .:.  :. ..:.:.
merascllllllplvhvsattpepceld___dedfrcvcnfsepqpd_wseafqcvsa
```

L6 ANSWER 3 OF 25 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-W05316 Protein DGENE
TITLE: Recombinant DNA encoding myelomonocytic differentiation antigen CD14 - used for producing recombinant CD14 for use in e.g. diagnosis of myeloid disorders such as leukaemia
INVENTOR: Goyert S M
PATENT ASSIGNEE: (GOYE-I)GOYERT S M
PATENT INFO: US 5543303 A 960806 11 pp
APPLICATION INFO: US 88-276794 881128
PRIORITY INFO: US 88-276794 881128
US 90-536163 900608
US 91-787763 911106
US 92-916806 920722
US 93-165583 931213
PAT. SEQ. LOC: Claim 1; Fig 3
DATA ENTRY DATE: 03 JAN 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-370638 [37]
CROSS REFERENCES: N-PSDB: 96N-T39716; N-PSDB; 96N-T39717
DESCRIPTION: Myelomonocytic differentiation antigen CD14
KEYWORD: Myelomonocytic differentiation antigen; CD14; myeloid leukaemia; diagnosis
ORGANISM: Homo sapiens
ABSTRACT:

Human myelomonocytic differentiation antigen CD14 (W05316) is an antigen useful in the diagnosis of mature myeloid leukemia. Its amino acid sequence was deduced from a cDNA clone (T39717) obtd. by screening COS 7 cell transfectants with monoclonal antibodies to CD14. Large amts. of CD14 can be produced by expression in transformed host cells; mature, glycosylated CD14 is produced in mammalian host cells, and nonglycosylated CD14 in prokaryotic hosts

AMINO ACID COUNTS: 43 A; 22 R; 15 N; 16 D; 0 B; 11 C; 14 Q; 20 E; 0 Z; 24 G; 7 H; 4 I; 62 L; 9 K; 6 M; 10 F; 30 P; 29 S; 18 T; 5 W; 2 Y; 28 V;
SEQUENCE LENGTH: 375
SEQUENCE

```
      1 merascllll llplvhvsat tpepceldde dfrcvcnfse pqpdwseafq
     51 cvsaveveih agglnlepfl krvdadadpr qyadtvkalr vrrltvgaaq
    101 vpaqllvgal rvlaysrlke ltledlkitg tmpplpleat glalsslrlr
    151 nvswatgrsw laelqqwlkp glkvlsiaqa hspafsceqv rafpaltsld
    201 lsdnpglger glmaalcphk fpaiqnlalr ntgmetptgv caalaaagvq
    251 phsldlshns lratvnpsap rcmwssalns lnlsfagleq vpkglpaklr
    301 vldlscnrln rapqpdelpe vdnltldgnp flvpgtalph egsmnsgvvp
    351 acarstlsvg vsgtlvllqg argfa
```

FEATURE TABLE:

Key	Location	Qualifier	
Peptide	1..19	label	Sig_peptide
Modified_site	37..39	label	Glycosylation
		note	"potential N-linked glycosylation site"
Modified_site	151..153	label	Glycosylation
		note	"potential N-linked

```
                |         |          |glycosylation site"
Modified_site   |266..268|label     |Glycosylation
                |         |note      |"potential N-linked
                |         |          |glycosylation site"
Modified_site   |282..284|label     |Glycosylation
                |         |note      |"potential N-linked
                |         |          |glycosylation site"
Modified_site   |323..325|label     |Glycosylation
                |         |note      |"potential N-linked
                |         |          |glycosylation site"

ALIGN Smith-Waterman score:  94
      58 aa overlap starting at 1
      mgpsscllliliplqlinpgstqcsldsvmdkkikdvlnsleyspspiskklscasv
      :  .:::::.:.::...       : ::   :. .. : :  : .:.  :. ..:.:.
      merascllllllplvhvsattpepceld___dedfrcvcnfsepqpd_wseafqcvsa
```

L6 ANSWER 4 OF 25 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R71730 Protein DGENE
TITLE: New merosin fragments, corresp. DNA and antibodies - for diagnosing tumour malignancy, promoting or inhibiting neurite growth and promoting cell attachment
INVENTOR: Engvall E; Leivo I
PATENT ASSIGNEE: (LJOL-N)LA JOLLA CANCER RES FOUND
PATENT INFO: WO 9508628 A2 950330 65 pp
APPLICATION INFO: WO 94-US10730 940921
PRIORITY INFO: US 93-125077 930922
PAT. SEQ. LOC: Claim 5; Fig 6
DATA ENTRY DATE: 01 MAY 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 95-139597 [18]
CROSS REFERENCES: N-PSDB: 95N-Q86480 AND T17419
DESCRIPTION: Merosin major subunit
KEYWORD: Human; 380-400 kD; merosin; major subunit; placenta; straited muscle; peripheral nerve; trophoblast; Schwann cell neoplasm; 65 kD subunit; 80 kD subunit; merosin polypeptide; merosin subunit; M chain; laminin M chain; antigen; antibody; detection; tumour; malignancy; neurite outgrowth; inhibitor; cell attachment
ORGANISM: Homo sapiens
ABSTRACT:

This sequence represents the human 380-400 kD merosin major subunit. Merosin is an isoform of laminin and shows structural and sequence similarity to the human laminin A chain. Mature human merosin is 30 amino acids larger than the human laminin A chain. Similarly to all laminin chains, the merosin protein has distinct domains which are predicted to have globular regions, cysteine-rich rod-like regions and helical structures (see features table). Merosin has a large globular domain at the carboxy terminal end. The merosin protein has an apparent mol. wt. of about 800 kD and is composed of four polypeptides with molecular weights of 300, 200, 200 and 80 kD. The 300 kD polypeptide is joined to the 200 kD polypeptides by disulphide bonds and the 300 and 80 kD polypeptides comprise the major subunit protein given in R71729. Merosin is found in placenta, straited muscle, peripheral nerve, trophoblasts and human Schwann cell neoplasms, amoung other tissues. The 380-400 major merosin subunit also yields a 65 kD subunit. The 380-400 merosin subunit has been designated merosin polypeptide, merosin subunit, M chain or laminin M chain. Fragments of the merosin protein may be used as antigens to raise anti-merosin antibodies. These antibodies may be used in the detection of merosin, as the absence of merosin in a tumour sample indicates malignancy. Contacting a neurone with merosin promotes neurite outgrowth. The merosin polypeptide may also be used in contacting inhibitors of neurite outgrowth, thereby also promoting the outgrowth. Merosin also promotes cell attachment. The merosin gene has been mapped to chromosome 6, more precisely to bands 6q22->q23

AMINO ACID COUNTS:205A; 159R; 162N; 183D; 0 B; 162C; 119Q; 202E; 0 Z; 261G; 71 H; 166I; 246L; 184K; 46 M; 103F; 173P; 194S; 193T; 29 W; 96 Y; 156V;
SEQUENCE LENGTH: 3110

SEQUENCE

```
       1 mpgaagvlll lllsgglggv qaqrpqqqrq sqahqqrglf pavlnlasna
      51 littnatcge kgpemycklv ehvpgqpvrn pqcricnqns snpnqrhpit
     101 naidgkntww qspsiknqie yhyvtitldl qqvfqiayvi vkaansprpg
     151 nwilersldd veykpwqyha vtdtecltly niyprtgpps yakddevict
     201 sfyskihple ngeihislin grpsaddpsp elleftsary irlrfqrirt
     251 lnadlmmfah kdpreidpiv trryyysvkd isvggmcicy gharacpldp
     301 atnksrcece hntcgdscdq ccpgfhqkpw ragtfltkte ceacnchgka
     351 eecyydenva rrnlslnirg kyigggvcin ctqntaginc etctdgffrp
     401 kgvspnyprp cqpchcdpig slnevcvkde kharrglapg schcktgfgg
     451 vscdrcargy tgypdckacn csglgskned pcfgpcicke nveggdcsrc
     501 ksgffnlqed nwkgcdecfc sgvsnrcqss ywtygkiqdm sgwyltdlpg
     551 rirvapqqdd ldspqqisis naearqalph syywsapapy lgnklpavgg
     601 qltftisydl eeeeedterv lqlmiilegn dlsistaqde vylhpseeht
     651 nvlllkeesf tihgthfpvr rkefmtvlan lkrvllqity sfgmdaifrl
     701 ssvnlesavs yptdgsiaaa vevcqcppgy tgsscescwp rhrrvngtif
     751 ggicepcqcf ghaescddvt geclnckdht ggpycdkclp gfygeptkgt
     801 sedcqpcacp lnipsnnfsp tchldrslgl icdgcpvgyt gprcercaeg
     851 yfgqpsvpgg scqpcqcndn ldfsipgscd slsgsclick pgttgrycel
     901 cadgyfgdav dakncqpcrc naggsfsevc hsqtgqcecr anvqgqrcdk
     951 ckagtfglqs argcvpcncn sfgsksfdce esgqcwcqpg vtgkkcdrca
    1001 hgyfnfqegg ctacecshlg nncdpktgrc icppntigek cskcapntwg
    1051 hsittgckac ncstvgsldf qcnvntgqcn chpkfsgakc tecsrghwny
    1101 prcnlcdcfl pgtdattcds etkkcscsdq tgqctckvnv egihcdrcrp
    1151 gkfgldaknp lgcsscycfg tttqcseakg lirtwvtlka eqtilplvde
    1201 alqhtttkgi vfqhpeivah mdlmredlhl epfywklpeq fegkklmayg
    1251 gklkyaiyfe areetgfsty npqviirggt pthariivrh maapligqlt
    1301 rheiemteke wkyygddprv hrtvtredfl dilydihyil ikatygnfmr
    1351 qsriseisme vaeqgrgttm tppadliekc dcplgysgls ceaclpgfyr
    1401 lrsqpggrtp gptlgtcvpc qcnghsslcd petsicqncq hhtagdfcer
    1451 calgyygivk glpndcqqca cplisssnnf spscvaegld dyrctacprg
    1501 yegqycerca pgytgspgnp ggscqececd pygslpvpcd pvtgfctcrp
    1551 gatgrkcdgc khwharegwe cvfcgdectg lllgdlarle qmvmsinltg
    1601 plpapykmly glenmtqelk hllspqrape rliqlaegnl ntlvtemnel
    1651 ltratkvtad geqtgqdaer tntrakslge fikelardae avnekaikln
    1701 etlgtrdeaf ernleglqke idqmikelrr knletqkeia edelvaaeal
    1751 lkkvkklfge srgeneemek dlrekladyk nkvddawdll reatdkirea
    1801 nrlfavnqkn mtalekkkea vesgkrqien tlkegndild eanrladein
    1851 siidyvediq tklppmseel ndkiddlsqe ikdrklaekv sqaeshaaql
    1901 ndssavldgi ldeaknisfn ataafkaysn ikdyideaek vakeakdlah
    1951 eatklatgpr gllkedakgc lqksfrilne akklandvke nedhlnglkt
    2001 rienadarng dllrtlndtl gklsaipndt aaklqavkdk arqandtakd
    2051 vlaqitelhq nldglkknyn kladsvaktn avvkdpsknk iiadadatvk
    2101 nleqeadrli dklkpikele dnlkknisei kelinqarkq ansikvsvss
    2151 ggdcirtykp eikkgsynni vvnvktavad nllfylgsak fidflaiemr
    2201 kgkvsflwdv gsgvgrveyp dltiddsywy rivasrtgrn gtisvraldg
    2251 pkasivpsth hstsppgyti ldvdanamlf vggltgklkk adavrvitft
    2301 gcmgetyfdn kpiglwnfre kegdckgctv spqvedsegt iqfdgegyal
    2351 vsrpirwypn istvmfkfrt fsssallmyl atrdlrdfms veltdghikv
    2401 sydlgsgmas vvsnqnhndg kwksftlsri qkqanisivd idtnqeenia
    2451 tsssgnnfgl dlkaddkiyf gglptlrnls mkarpevnlk kysgclkdie
    2501 isrtpynils spdyvgvtkg cslenvytvs fpkpgfvels pvpidvgtei
    2551 nlsfstknes giillgsggt papprrkrrq tgqayyvill nrgrlevhls
    2601 tgartmrkiv irpepnlfhd grehsvhver trgiftvqvd enrrymqnlt
    2651 veqpievkkl fvggappefq psplrnippf egciwnlvin svpmdfarpv
    2701 sfknadigrc ahqklreded gaapaeiviq pepvptpafp tptpvlthgp
    2751 caaesepall igskqfglsr nshiaiafdd tkvknrltie levrteaesg
    2801 llfymaainh adfatvqlrn glpyfsydlg sgdthtmipt kindgqwhki
```

```
2851 kimrskqegi lyvdgasnrt ispkkadild vvgmlyvggl pinyttrrig
2901 pvtysidgcv rnlhmaeapa dleqptssfh vgtcfanaqr gtyfdgtgfa
2951 kavggfkvgl dllvefefat ttttgvllgi ssqkmdgmgi emideklmfh
3001 vdngagrfta vydagvpghl cdgqwhkvta nkikhrielt vdgnqveaqs
3051 pnpastsadt ndpvfvggfp ddlkqfgltt sipfrgcirs lkltkgtash
3101 wrlilprpwn
```

FEATURE TABLE:

Key	Location	Qualifier	
Region	1..2298	note	"region encoded by Q86480"
Peptide	1..22	note	"Signal peptide"
Domain	23..286	label	Domain VI
		note	"predicted to form globular
			structure"
Modified_site	55..57	note	"N-linked glycosylation site"
Modified_site	89..91	note	"N-linked glycosylation site"
Domain	287..527	label	Domain V
		note	"contains four and one half
			Cystein-rich EGF-like repeats,
			predicted to have rigid
			rod-like structure"
Modified_site	303..305	note	"N-linked glycosylation site"
Modified_site	363..365	note	"N-linked glycosylation site"
Modified_site	380..382	note	"N-linked glycosylation site"
Modified_site	470..472	note	"N-linked glycosylation site"
Domain	528..723	label	Domain IVb
		note	"predicted to form globular
			structure"
Domain	724..1175	label	Domain IIIb
		note	"contains ten and one half
			Cystein-rich EGF-like repeats,
			predicted to have rigid
			rod-like structure"
Modified_site	746..748	note	"N-linked glycosylation site"
Modified_site	1061..1063	note	"N-linked glycosylation site"
Domain	1176..1379	label	Domain IVa
		note	"predicted to form globular
			structure"
Domain	1180..1573	label	Domain IIIa
		note	"contains four Cystein-rich
			EGF-like repeats, predicted
			to have rigid rod-like
			structure"
Domain	1574..2153	label	Domain I+II
		note	"forms two B-type chains,
			forms triple coiled-coil
			structure"
Modified_site	1597..1599	note	"N-linked glycosylation site"
Modified_site	1614..1616	note	"N-linked glycosylation site"
Modified_site	1700..1702	note	"N-linked glycosylation site"
Modified_site	1810..1812	note	"N-linked glycosylation site"
Modified_site	1901..1903	note	"N-linked glycosylation site"
Modified_site	1916..1918	note	"N-linked glycosylation site"
Modified_site	1920..1922	note	"N-linked glycosylation site"
Modified_site	2017..2019	note	"N-linked glycosylation site"

```
Modified_site  |2028..2030|note     |"N-linked glycosylation site"
Modified_site  |2045..2047|note     |"N-linked glycosylation site"
Modified_site  |2126..2128|note     |"N-linked glycosylation site"
Domain         |2154..3110|label    |Domain G
               |          |note     |"forms large globule at end of
               |          |         |laminin long arm"
Modified_site  |2240..2242|note     |"N-linked glycosylation site"
Modified_site  |2360..2362|note     |"N-linked glycosylation site"
Modified_site  |2435..2437|note     |"N-linked glycosylation site"
Modified_site  |2478..2480|note     |"N-linked glycosylation site"
Modified_site  |2551..2553|note     |"N-linked glycosylation site"
Modified_site  |2558..2560|note     |"N-linked glycosylation site"
Modified_site  |2648..2650|note     |"N-linked glycosylation site"
Modified_site  |2868..2870|note     |"N-linked glycosylation site"
   Modified_site  |2893..2895|note     |"N-linked glycosylation site"

ALIGN Smith-Waterman score:  97
      64 aa overlap starting at 939
      c_asvksqgrpsscpag____mavtgca___c_gygcgswdvqlettchcqcsvvdwtta
      : :.:..: : ..: ::    ... ::.   : ..:  :.: .    : :: .:.
      cranvqgq_rcdkckagtfglqsargcvpcncnsfgsksfdceesgqcwcqpgvtgkkcd
      rcch
      :: :
      rcah
```

L6 ANSWER 5 OF 25 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 90P-R07450 protein DGENE
TITLE: DNA encoding TNF binding protein and TNF- receptor - used in tumour treatment and to understand mechanismsm to TNF action
INVENTOR: Hauptmann R; Himmler A; Maurer-Fogy I; Stratowa C
PATENT ASSIGNEE: (BOEH)BOEHRINGER INGELHEIMINT
PATENT INFO: EP 393438 A 901024 51 pp
APPLICATION INFO: EP 90-106624 900406
PRIORITY INFO: DE 89-3920282 890621
DE 89-3913101 890421
PAT. SEQ. LOC: Disclosure; Fig 8(1-2)
DATA ENTRY DATE: 29 JAN 1991 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: German
OTHER SOURCE: 90-321987 [43]
CROSS REFERENCES: N-PSDB: 90N-Q06284
DESCRIPTION: Rat Tumour Necrosis Factor-Receptor from raTNF-R8 cDNA
KEYWORD: Tumour necrosis factor binding protein; TNF-BP; TNF-receptor; raTNF-R8
ORGANISM: Rat rattus
ABSTRACT:

A rat brain cDNA analogue of the HS913T cDNA library from rat glioma cell line C6 (ATCC CCL107) is prepared in lambda-gt11. The isolated clone raTNF-R8 is used as probe to isolated the entire human TNF receptor, as represented in Q06285. See also Q06282-Q06285

AMINO ACID COUNTS: 22 A; 26 R; 20 N; 16 D; 0 B; 31 C; 16 Q; 25 E; 0 Z; 28 G; 14 H; 16 I; 48 L; 20 K; 9 M; 17 F; 44 P; 31 S; 29 T; 5 W; 8 Y; 36 V;
SEQUENCE LENGTH: 461
SEQUENCE

```
  1 mglpivpgll lslvllallm gihpsgvtgl vpslgdrekr dnlcpqgkya
 51 hpknnsicct kchkgtylvs dcpspgqetv celshkgtft asqnhvrqcl
101 scktcrkemf qveispckad mdtvcgckkn qfqrylseth fqcvdcspcf
151 ngtvtipcke kqntvcncha gfflsgnect pcshckknqe cmklclppva
201 nvtnpqdsgt avllplvifl glcllffici sllcrypqwr prvysiicrd
251 sapvkevege givtkpltpa sipafsanpg fnptlgfstt prfshpvsst
301 pispvfgpsn whnfvppvre vvptqgadpl lygclnpvpi papvrkwedv
351 vaaqpqrldt adpamlyavv dgvpptrwke fmrllglseh eierlelqng
401 rclreahysm leawrrrtpr deatldvvgr vlcdmnlrgc leniretles
451 pahsstthlp r
```

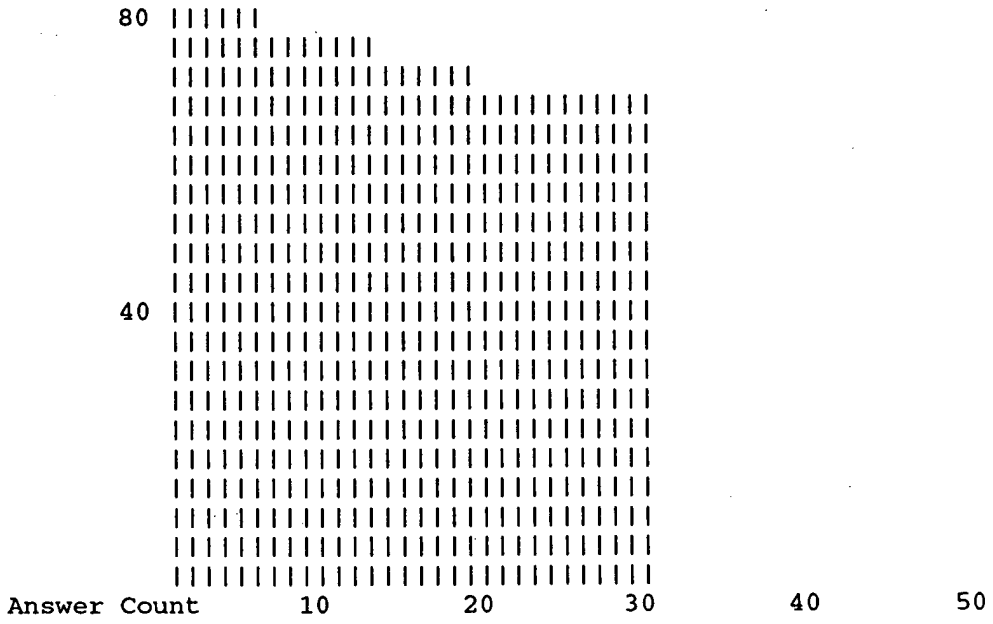
ALIGN Smith-Waterman score: 89
99 aa overlap starting at 7

```
pssclllilipllqlinpgstqcsldsvmdkkikdvlnsleyspspiskklscasv_ksq
:.  : :.:. ::. :.:...   . :. :.. .: :       : .... :..  :.
pgllslvllallmgihpsgvtglvpslgdrekrdnlcpqgkyahpknnsicctkchkgt
grpsscpagmavtgcacgygcgswdvqlettchc_qcsv
   :.::.    : :  ..  :.. .. . . .: .:..
ylvsdcpspgqetvcelshk_gtftasqnhvrqclsckt
```

Mouse C18

MKPTLCFLFILVSLFPLIVPGNAQCSFESLVDQRIKEALSRQEPKTISCTSVTSSGRLASCPAGMVVTGCACGYGCGS
WDIRNGNTCHCQCSVMDWASARCCRMA

30 ANSWERS FOUND ABOVE A THRESHOLD OF 65



L8 ANSWER 1 OF 30 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W44299 Protein DGENE
TITLE: Human serrate-2 gene expression products - used to regulate stem cell differentiation, useful in treating neoplasms, e.g. leukaemia
INVENTOR: Itoh A; Sakano S
PATENT ASSIGNEE: (ASAH)ASAHI KASEI KOGYO KK
PATENT INFO: WO 9802458 A1 980122 103 pp
APPLICATION INFO: WO 97-JP2414 970711
PRIORITY INFO: JP 97-124063 970514
JP 96-186220 960716
PAT. SEQ. LOC: Claim 3; Page 62-68
DATA ENTRY DATE: 19 JUN 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: Japanese
OTHER SOURCE: 98-110528 [10]
CROSS REFERENCES: N-PSDB: 98N-V15181
DESCRIPTION: Human serrate 2
KEYWORD: Human; serrate 2; regulation; stem cell; differentiation; neoplasm; leukaemia; endothelial cell; tumour
ORGANISM: Homo sapiens
ABSTRACT:

The present sequence represents human serrate 2. The present invention also describes a method for the preparation of the polypeptides, and antibodies binding to the polypeptide and its fragments. The polypeptide and its fragments expressed by the serrate-2 gene can be used to inhibit stem (especially blood stem) cell differentiation and to inhibit endothelial cell growth. They may be incorporated in a cell culture media for culturing undifferentiated stem cells. They can also be used for treatment of neoplasms such as leukaemia. The antibodies can be used for the diagnosis of malignant tumours

AMINO ACID COUNTS:87 A; 76 R; 59 N; 76 D; 0 B; 130C; 31 Q; 70 E; 0 Z; 143G; 34 H; 30 I; 72 L; 33 K; 5 M; 35 F; 88 P; 75 S; 50 T; 26 W; 30 Y; 62 V;
SEQUENCE LENGTH: 1212
SEQUENCE

```
   1 mgyfelqlsa lrnvngells gaccdgdgrt traggcghde cdtyvrvclk
  51 eyqakvtptg pcsyghgatp vlggnsfylp pagaagdrar araraggdqd
 101 pglvvipfqf awprsftliv eawdwdndtt pneellierv shagminped
 151 rwkslhfsgh vahlelqirv rcdenyysat cnkfcrprnd ffghytcdqy
 201 gnkacmdgwm gkeckeavck qgcnllhggc tvpgecrcsy gwqgrfcdec
 251 vpypgcvhgs cvepwqcnce tnwggllcdk dlnycgshhp ctnggtcina
 301 epdqyrctcp dgysgrncek aehactsnpc anggschevp sgfechcpsg
 351 wsgptcaldi decasnpcaa ggtcvdqvdg fecicpeqwv gatcqldane
 401 cegkpclnaf scknliggyy cdcipgwkgi nchinvndcr gqcqhggtck
 451 dlvngyqcvc prgfggrhce lerdkcassp chsgglcedl adgfhchcpq
 501 gfsgplcevd vdlcepspcr ngarcynleg dyycacpddf ggkncsvpre
 551 pcpggacrvi dgcgsdagpg mpgtaasgvc gphgrcvsqp ggnfscicds
 601 gftgtychen iddclgqpcr nggtcidevd afrcfcpsgw egelcdtnpn
 651 dclpdpchsr grcydlvndf ycacddgwkg ktchsrefqc daytcsnggt
 701 cydsgdtfrc acppgwkgst cavaknsscl pnpcvnggtc vgsgasfsci
 751 crdgwegrtc thntndcnpl pcynggicvd gvnwfrceca pgfagpdcri
 801 nidecqsspc aygatcvdei ngyrcscppg ragprcqevi gfgrscwsrg
 851 tpfphgsswv edcnscrcld grrdcskvwc gwkpcllagq pealsaqcpl
 901 gqrclekapg qclrppceaw gecgaeepps tpclprsghl dnncarltlh
 951 fnrdhvpqgt tvgaicsgir slpatravar drllvllcdr assgasalev
```

```
  1001 avsfspardl pdssliqgaa haivaaitqr gnsslllavt evkvetvvtg
  1051 gsstgllvpv lcgafsvlwl acvvlcvwwt rkrrkerers rlpreesann
  1101 qwaplnpirn pierpgghkd vlyqcknftp pprradealp gpaghaavre
  1151 deededlgrg eedsleaekf lshkftkdpg rspgrpahwa sgpkvdnrav
  1201 rsinearyag ke
ALIGN Smith-Waterman score:  94
   105 aa overlap starting at 253
   fplivpgn_a___qcsfes_____lvdqrikealsrqepkti_sctsvtssgrlascpag
   .:  : :. .   ::. :.     : :. ..   :~.   .  .: ..  .    .:: :
   ypgcvhgscvepwqcncetnwgg1lcdkdlnycgshhpctnggtcinaepdqyrctcpdg
   mvvtgcacg_ygcgswdirngntch_______cqcsvmdwasarc
   .   .:  . ..: :    ::..::       :.:   :..  :
   ysgrncekaehactsnpcanggschevpsgfechcp_sgwsgptc
```

L8 ANSWER 2 OF 30 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W44298 Protein DGENE
TITLE: Human serrate-2 gene expression products - used to regulate stem cell differentiation, useful in treating neoplasms, e.g. leukaemia
INVENTOR: Itoh A; Sakano S
PATENT ASSIGNEE: (ASAH)ASAHI KASEI KOGYO KK
PATENT INFO: WO 9802458 A1 980122 103 pp
APPLICATION INFO: WO 97-JP2414 970711
PRIORITY INFO: JP 97-124063 970514
JP 96-186220 960716
PAT. SEQ. LOC: Claim 2; Page 57-62
DATA ENTRY DATE: 19 JUN 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: Japanese
OTHER SOURCE: 98-110528 [10]
CROSS REFERENCES: N-PSDB: 98N-V15181
DESCRIPTION: Human serrate 2 protein fragment
KEYWORD: Human; serrate 2; regulation; stem cell; differentiation; neoplasm; leukaemia; endothelial cell; tumour
ORGANISM: Homo sapiens
ABSTRACT:

The present sequence represents a human serrate 2 protein fragment. The present invention also describes a method for the preparation of the polypeptides, and antibodies binding to the polypeptide and its fragments. The polypeptide and its fragments expressed by the serrate-2 gene can be used to inhibit stem (especially blood stem) cell differentiation and to inhibit endothelial cell growth. They may be incorporated in a cell culture media for culturing undifferentiated stem cells. They can also be used for treatment of neoplasms such as leukaemia. The antibodies can be used for the diagnosis of malignant tumours

AMINO ACID COUNTS:72 A; 58 R; 52 N; 68 D; 0 B; 126C; 29 Q; 54 E; 0 Z; 132G; 30 H; 27 I; 59 L; 24 K; 5 M; 31 F; 73 P; 67 S; 47 T; 21 W; 28 Y; 52 V;
SEQUENCE LENGTH: 1055
SEQUENCE

```
   1 mgyfelqlsa lrnvngells gaccdgdgrt traggcghde cdtyvrvclk
  51 eyqakvtptg pcsyghgatp vlggnsfylp pagaagdrar araraggdqd
 101 pglvvipfqf awprsftliv eawdwdndtt pneellierv shagminped
 151 rwkslhfsgh vahlelqirv rcdenyysat cnkfcrprnd ffghytcdqy
 201 gnkacmdgwm gkeckeavck qgcnllhggc tvpgecrcsy gwqgrfcdec
 251 vpypgcvhgs cvepwqcnce tnwgqllcdk dlnycgshhp ctnggtcina
 301 epdqyrctcp dgysgrncek aehactsnpc anggschevp sgfechcpsg
 351 wsgptcaldi decasnpcaa ggtcvdqvdg fecicpeqwv gatcqldane
 401 cegkpclnaf scknliggyy cdcipgwkgi nchinvndcr gqcqhggtck
 451 dlvngyqcvc prgfggrhce lerdkcassp chsgglcedl adgfhchcpq
 501 gfsgplcevd vdlcepspcr ngarcynleg dyycacpddf ggkncsvpre
 551 pcpggacrvi dgcgsdagpg mpgtaasgvc gphgrcvsqp ggnfscicds
 601 gftgtychen iddclgqpcr nggtcidevd afrcfcpsgw egelcdtnpn
 651 dclpdpchsr grcydlvndf ycacddgwkg ktchsrefqc daytcsnggt
 701 cydsgdtfrc acppgwkgst cavaknsscl pnpcvnggtc vgsgasfsci
 751 crdgwegrtc thntndcnpl pcynggicvd gvnwfrceca pgfagpdcri
 801 nidecqsspc aygatcvdei ngyrcscppg ragprcqevi gfgrscwsrg
 851 tpfphgsswv edcnscrcld grrdcskvwc gwkpcllagq pealsaqcpl
 901 gqrclekapg qclrppceaw gecgaeepps tpclprsghl dnncarltlh
 951 fnrdhvpqgt tvgaicsgir slpatravar drllvllcdr assgasalev
```

```
     1001 avsfspardl pdssliqgaa haivaaitqr gnsslllavt evkvetvvtg
     1051 gsstg
ALIGN Smith-Waterman score:  94
      105 aa overlap starting at 253
      fplivpgn_a___qcsfes_____lvdqrikealsrqepkti_sctsvtssgrlascpag
      .:  : :. .   ::. :.     : :. ..   :..   . .: ..  .    .:: :
      ypgcvhgscvepwqcncetnwgg1lcdkdlnycgshhpctnggtcinaepdqyrctcpdg
      mvvtgcacg_ygcgswdirngntch_____cqcsvmdwasarc
      .   .:  . ..: :    ::..::       :.:-   :..  :
      ysgrncekaehactsnpcanggschevpsgfechcp_sgwsgptc
```

L8 ANSWER 3 OF 30 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-W05834 Protein DGENE
TITLE: Vertebrate Serrate protein and related DNA - used to treat or prevent malignancies characterised by increased Notch activity
INVENTOR: Artavanis-Tsakonas S; Gray G E; Henrique D M P; Ish-Horowicz D; Lewis J H; Mann R S; Myat A M
PATENT ASSIGNEE: (IMCR) IMPERIAL CANCER RES TECHNOLOGY
(UYYA) UNIV YALE
PATENT INFO: WO 9627610 A1 960912 161 pp
APPLICATION INFO: WO 96-US3172 960307
PRIORITY INFO: US 95-400159 950307
PAT. SEQ. LOC: Claim 5; Page 104-107
DATA ENTRY DATE: 28 JAN 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-425379 [42]
CROSS REFERENCES: N-PSDB: 96N-W05834
DESCRIPTION: Human Serrate-2 (HJ2)
KEYWORD: Serrate-2; human jagged-2; HJ2; Notch; cell differentiation; cell fate; central nervous system; cancer; tissue repair; therapy; diagnosis; antibody
ORGANISM: Homo sapiens
ABSTRACT:

Human Serrate-1 (W05833) and human Serrate-2 (W05833) are ligands for the zygotic neurogenic locus Notch, and are believed to play a major role in determining cell fates (differentiation) in the central nervous system. Their amino acid sequences were deduced from cDNA clones (see also T40090-91) isolated from human foetal brain cDNA libraries. The proteins, antibodies raised to them, and encoding nucleic acids can be used in the detection of Serrate sequences and in the treatment of disorders of cell fate or differentiation, partic. cancer, nervous system disorders and in tissue repair or regeneration

AMINO ACID COUNTS: 85 A; 87 R; 58 N; 64 D; 0 B; 137C; 34 Q; 56 E; 0 Z; 149G; 34 H; 32 I; 72 L; 36 K; 10 M; 39 F; 90 P; 86 S; 59 T; 30 W; 28 Y; 71 V;
SEQUENCE LENGTH: 1257
SEQUENCE

```
   1 minpedrwks lhfsghvahl elqirvrcde nyysatcnkf crprndffgh
  51 ytcdqygnka cmdgwmgkec keavckqgcn llhggctvpg ecrcsygwqg
 101 rfcdecvpyp gcvhgscvep wqcncetnwg gllcdkdlny cgshhpctng
 151 gtcinaepdq yrctcpdgys grncekaeha ctsnpcangg schevpsgfe
 201 chcpsgwsgp tcaldideca snpcaaggtc vdqvdgfeci cpeqwvgatc
 251 qldanecegk pclnafsckn liggyycdci pgwkginchi nvndcrgqcq
 301 hggtckdlvn gyqcvcprgf ggrhcelerd kcasspchsg glcedladgf
 351 hchcpqgfsg plcevdvdlc epspcrngar cynlegdyyc acpddfggkn
 401 csvprepcpg gacrvidgcg sdagpgmpgt aasgvcgphg rcvsqpggnf
 451 scicdsgftg tychenlddc lgqpcrnggt cidevdafrc fcpsgwegel
 501 cdtnpndclp dpchsrgrcy dlvndfycac ddgwkgktch srefqcdayt
 551 csnggtcyds gdtfrcacpp gwkgstcava knssclpnpc vnggtcvgsg
 601 asfscicrdg wegrtcthnt ndcnplpcyn ggicvdgvnw frcecapgfa
 651 gpdcrinide cqsspcayga tcvdeingyr cscppgragp rcqevigfgr
 701 scwsrgtpfp hgsswvedcn scrcldgrrd cskvwcgwkp cllagqpeal
 751 saqcplgqrc lekapgqclr ppceawgecg aeeppstpcl prsghldnnc
 801 arltlhfnrd hvpqgttvga icsgirslpa travardrll vllcdrassg
 851 asavevavsf spardlpdss liqgaahaiv aaitqrgnss lllavtevkv
 901 etvvtggsst gllvpvlcga fsvlwlacvv lcvwwtrkrr kerersrlpr
```

SEARCHED ON 26 OCT 1998
FILE LAST UPDATED: 18 OCT 1998 <19981018/UP>

```
 951 eesannqwap lnpirnpier pgghkdvlyq cknftppprr rcpgrpatrp
1001 sgrmrrtril aavrrtpwrr rssshtnspk ilaarrggrp tgpqapkwtt
1051 arsgasmrpa tsarevgrlq lgrdpgpsvg ampsagpggr ghvhsffilc
1101 kkttknknqm fifyvsltly klfsncqaen ngvfsdscyf ckvavrgtrc
1151 mkgeskgclr rhqivafvtr gcalftessf ysslgflcap gqsagethgc
1201 vgvahgcwwd pwlmvwpvav ggtrgcqwdl wlsvgptvvg gtlvidvala
1251 agtargc
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	1..912	label	Extracellular_domain
		note	"a deletion in the encoding cDNA clone results in loss of part of the Serrate-2 signal peptide and beginning of the DSL domain
Domain	26..70	label	DSL
		note	"region of homology with Drosophila Delta and Serrate, predicted to mediate binding with Notch"
Domain	75..735	label	ELR
		note	"epidermal growth factor-like repeat domain"
Region	75..105	label	ELR1
Region	106..140	label	ELR2
Region	141..180	label	ELR3
Region	181..218	label	ELR4
Region	219..256	label	ELR5
Region	257..294	label	ELR6
Region	295..331	label	ELR7
Region	332..369	label	ELR8
Region	370..407	label	ELR9
Region	408..435	label	Partial_ELR
Region	436..469	label	Partial_ELR
Region	470..507	label	ELR10
Region	508..545	label	ELR11
Region	546..584	label	ELR12
Region	585..622	label	ELR13
Region	623..660	label	ELR14
Region	664..701	label	ELR15
Region	702..718	label	Partial_ELR
Region	719..735	label	Partial_ELR
Domain	913..933	label	Transmembrane_domain
Domain	934..1257	label	Intracellular_domain

```
   ALIGN Smith-Waterman score:  94
      105 aa overlap starting at 109
      fplivpgn_a___qcsfes_____lvdqrikealsrqepkti_sctsvtssgrlascpag
      .:  : :. .   ::. :.     : :. ..   :..   .  .: ..  .    ::: :
      ypgcvhgscvepwqcncetnwggllcdkdlnycgshhpctnggtcinaepdqyrctcpdg
      mvvtgcacg_ygcgswdirngntch_______cqcsvmdwasarc
      .   .:  . ..: :    ::..::       :.:    :..  :
      ysgrncekaehactsnpcanggschevpsgfechcp_sgwsgptc
```

L8 ANSWER 4 OF 30 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W54234 peptide DGENE
TITLE: Detection and therapy of cervical cancer - using specific cervical cancer-associated proteins as targets for treatment or as indicators for detection
INVENTOR: Keesee S K; Obar R; Wu Y
PATENT ASSIGNEE: (MATR-N)MATRITECH INC
PATENT INFO: WO 9809170 A2 980305 79 pp
APPLICATION INFO: WO 97-US14526 970819
PRIORITY INFO: US 96-705660 960830
PAT. SEQ. LOC: Claim 12; Page 55-56
DATA ENTRY DATE: 10 AUG 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 98-230271 [20]
DESCRIPTION: Human TDP-43 protein
KEYWORD: Cervical cancer-associated protein; CvC; tryptic peptide; human; detection; treatment; TDP-43; TAR DNA binding protein; non-chromatin protein
ORGANISM: Homo sapiens
ABSTRACT:

This protein is the human TDP-43 protein (also known as TAR DNA binding protein) which is used to obtain tryptic peptides which are used in a method for detecting cervical cancer. The method involves detecting the presence of a cervical cancer-associated protein (CvC) in a tissue or body fluid sample. The CvC is characterised as having a molecular weight of 44900-69400 Daltons as determined by sodium dodecyl-sulphate (SDS)-PAGE techniques and an isoelectric point (pI) of 5.1-6.6 as determined by standard isoelectric focusing techniques. The protein is further characterised as being a non-chromatin protein which is detectable at a higher level in a human cervical cancer cell than in a normal human cervical cell, as determined by 2D-gel electrophoresis. The methods can be used for the early and rapid detection of cervical cancer, for treating cervical cancers and for monitoring the efficacy of such treatment

AMINO ACID COUNTS:26 A; 20 R; 28 N; 22 D; 0 B; 6 C; 24 Q; 22 E; 0 Z; 55 G; 5 H; 14 I; 21 L; 20 K; 18 M; 22 F; 16 P; 41 S; 15 T; 6 W; 8 Y; 25 V;
SEQUENCE LENGTH: 414
SEQUENCE

```
   1 mseyirvted endepieips eddgtvllst vtaqfpgacg lryrnpvsqc
  51 mrgvrlvegi lhapdagwgn lvyvvnypkd nkrkmdetda ssavkvkrav
 101 qktsdlivlg lpwktteqdl keyfstfgev lmvqvkkdlk tghskgfgfv
 151 rfteyetqvk vmsqrhmidg rwcdcklpns kqsqdeplrs rkvfvgrcte
 201 dmtedelref fsqygdvmdv fipkpfrafa fvtfaddqia qslcgedlii
 251 kgisvhisna epkhnsnrql ersgrfggnp ggfgnqggfg nsrgggaglg
 301 nnqgsnmggg mnfgafsinp ammaaaqaal qsswgmmgml asqqnqsgps
 351 gnnqnqgnmq repnqafgsg nnsysgsnsg aaigwgsasn agsgsgfngg
 401 fgssmdskss gwgm
```

ALIGN Smith-Waterman score: 80
74 aa overlap starting at 224

```
kptlcflfilvslfplivpgnaqcsfeslvdqrikealsrqepktisctsvtssgrlasc
::   : :  :  :     ... :. :.:. . :.  .:  :::  :  ..  :::...
kpfrafaf__v_tfaddqiaqslcg_edliikgisvhisnaepkhnsnrqlersgrfggn
pagmvvtgcacgyg
:.:.   :   :.:
pggf___gnqggfg
```

SEARCHED ON 26 OCT 1998
FILE LAST UPDATED: 18 OCT 1998 <19981018/UP>

L8 ANSWER 5 OF 30 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 94P-R62046 protein DGENE
TITLE: Treating insulin-like growth factor related disease - by admin. of opt. modified binding protein, also new non-phosphorylated binding protein, e.g. for treating cancer or restenosis
INVENTOR: Cox G N; Russell D A
PATENT ASSIGNEE: (SYND)SYNERGEN INC
PATENT INFO: WO 9422466 A 941013 68 pp
APPLICATION INFO: WO 94-US3755 940406
PRIORITY INFO: US 93-45265 930407
PAT. SEQ. LOC: Disclosure; Page 10
DATA ENTRY DATE: 09 MAY 1995 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 94-332814 [41]
DESCRIPTION: Human mature IGFBP-1
KEYWORD: IGFBP-1; insulin-like growth factor binding protein-1; IGF-binding protein; BP-1; breast cancer; colon cancer; lung cancer; ovary cancer; liver cancer; osteosarcoma; glioma; rhabdomyosarcoma; restenosis; acromegaly; obesity; diabetic nephropathy; retinopathy
ORGANISM: Homo sapiens
ABSTRACT:

Muteins of human IGFBP-1 (mature protein given in R62046, signal peptide in R62047) have been produced containing a S98C or S101C substitution. Such muteins can be attached via thiol-reactive groups to PEG to form novel therapeutic agents

AMINO ACID COUNTS:26 A; 10 R; 9 N; 7 D; 0 B; 18 C; 9 Q; 22 E; 0 Z; 17 G; 6 H; 9 I; 17 L; 9 K; 3 M; 5 F; 17 P; 21 S; 9 T; 5 W; 6 Y; 9 V;
SEQUENCE LENGTH: 234
SEQUENCE

```
       1 apwqcapcsa eklalcppvs ascsevtrsa gcgccpmcal plgaacgvat
      51 arcarglscr alpgeqqplh altrgqgacv qesdasapha aeagspespe
     101 steiteeell dnfhlmapse edhsilwdai stydgskalh vtnikkwkep
     151 crielyrvve slakaqetsg eeiskfylpn cnkngfyhsr qcetsmdgea
     201 glcwcvyfwn gkripgspei rgdpncqiyf nvqn
```

ALIGN Smith-Waterman score: 87
37 aa overlap starting at 11

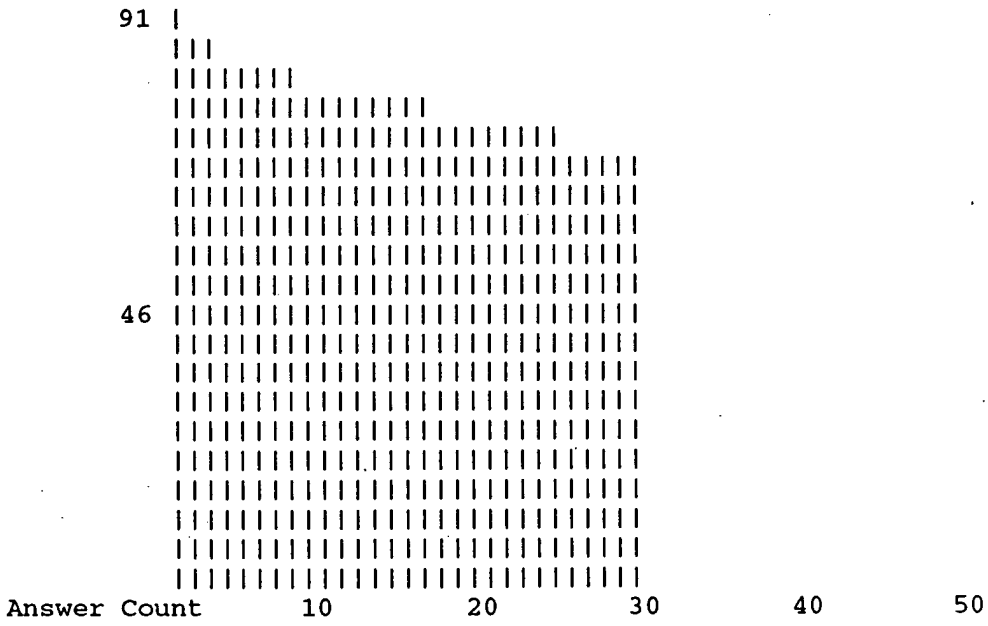
```
ealsrqepktisctsvtssgrlascpagmvvtgcacg
: :.   : . ::. :: :.  . ::   .  : :::
eklalcppvsascsevtrsagcgccpmcalplgaacg
```

REFERENCE: Derwent DGene Search Report
FRANZ-BACON, et al., USSN: 09/099,898
Atty. Docket No.: DX0744K

Mouse C19

MKNLSFPLLFLFFLVPELLGSSMPLCPIDEAIDKKIKQDFNSLFPNAIKNIGLNCWTVSSRGKLASCPEGTAVLSCSC
GSACGSWDIREEKVCHCQCARIDWTAARCCKLQVAS

29 ANSWERS FOUND ABOVE A THRESHOLD OF 68



L10 ANSWER 1 OF 29 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R74680 protein DGENE
TITLE: Genetically engineered tissue plasminogen activator - is modified at positions 44-50 and 296-302 and is non-glycosylated in K1 and K2 regions, has extended half-life and PAI-1 resistance
INVENTOR: Huang C; Huang P; Liu S
PATENT ASSIGNEE: (BIOE-N)BIOENGINEERING INST ACAD MILITARY
PATENT INFO: CN 1082111 A 940216 7 pp
APPLICATION INFO: CN 93-109234 930806
PRIORITY INFO: CN 93-109234 930806
PAT. SEQ. LOC: Claim 2; Fig 1 and Page 5
DATA ENTRY DATE: 04 JAN 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: Chinese
OTHER SOURCE: 95-162457 [22]
DESCRIPTION: t-PA mutein (N117Q, N184Q, delta 296-302, 44-50 replaced by KPIAEK)
KEYWORD: Tissue plasminogen activator; tPA; thrombolytic agent; mutein; deglycosylated kringle domain; PAI-1 resistance
ORGANISM: Synthetic
ABSTRACT:

The sequences given in R74678-R74689 are examples of preferred mutant versions of human tPA. In all the muteins, amino acids 296-302 of wild-type tPA (involved in interaction with PAI-1) have been deleted and the kringle domains have been deglycosylated by substn. of Asn 117 in K1 and Asn184 in K2 by Asp residues. Also, amino acids 44-50 of wild-type tPA are replaced by a motif which differs between different muteins. The modified tPA proteins have prolonged half-life, are resistant to PAI-1 and have affinity for fibrin; they are useful as thrombolytic agents

AMINO ACID COUNTS:33 A; 33 R; 20 N; 28 D; 0 B; 35 C; 28 Q; 27 E; 0 Z; 42 G; 14 H; 20 I; 39 L; 21 K; 5 M; 16 F; 28 P; 45 S; 25 T; 13 W; 24 Y; 23 V;
SEQUENCE LENGTH: 519
SEQUENCE

```
   1 syqvicrdek tqmiyqqhqs wlrpvlrsnr veycwcnsgr aqckpiaekc
  51 seprcfnggt cqqalyfsdf vcqcpegfag kcceidtrat cyedqgisyr
 101 gtwstaesga ectnwqssal aqkpysgrrp dairlglgnh nycrnpdrds
 151 kpwcyvfkag kyssefcstp acsegnsdcy fgqgsayrgt hsltesgasc
 201 lpwnsmilig kvytaqnpsa qalglgkhny crnpdgdakp wchvlknrrl
 251 tweycdvpsc stcglrqysq pqfrikgglf adiashpwqa aifaerflcg
 301 gilisscwil saahcfqerf pphhltvilg rtyrvvpgee eqkfevekyi
 351 vhkefdddty dndiallqlk sdssrcaqes svvrtvclpp adlqlpdwte
 401 celsgygkhe alspfyserl keahvrlyps srctsqhlln rtvtdnmlca
 451 gdtrsggpqa nlhdacqgds ggplvclndg rmtlvgiisw glgcgqkdvp
 501 gvytkvtnyl dwirdnmrp
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	1..49	label	finger_domain
		note	"amino acids 44-50 of F domain have been replaced by the sequence KPIAEK"
Domain	50..86	label	E_domain

```
                |         |note      |"growth factor domain"
Domain          |87..175  |label     |Kringle_1
                |         |note      |"substn. of Asn117 (corresp.
                |         |          |to position 116 in this
                |         |          |mutein) by Asp destroys an
                |         |          |N-linked glycosylation site"
Domain          |176..274 |label     |Kringle_2
                |         |note      |"substn. of Asn184 (corresp.
                |         |          |to position 183 in this
                |         |          |mutein) by Asp destroys an
                |         |          |N-linked glycosylation site"
Domain          |275..519 |label     |P_domain
                |         |note      |"amino acids 296-302 of native
                |         |          |tPA have been deleted; these
                |         |          |residues are involved in
                |         |          |interaction with PAI-1"
Disulfide_bond  |6..36    |          |
Disulfide_bond  |34..43   |          |
Disulfide_bond  |50..61   |          |
Disulfide_bond  |55..72   |          |
Disulfide_bond  |74..83   |          |
Disulfide_bond  |91..172  |          |
Disulfide_bond  |112..154 |          |
Disulfide_bond  |143..167 |          |
Disulfide_bond  |179..260 |          |
Disulfide_bond  |200..242 |          |
Disulfide_bond  |231..255 |          |
Disulfide_bond  |263..387 |          |
Disulfide_bond  |299..315 |          |
Disulfide_bond  |307..376 |          |
Disulfide_bond  |401..476 |          |
Disulfide_bond  |433..449 |          |
   Disulfide_bond |466..494|          |

ALIGN Smith-Waterman score:  91
      84 aa overlap starting at 8
      deaidkkikqdfnslfpnaikniglncwtvssrgklascpegtavlscscgsacgswdir
      ::  .   .:  . : :   .:    ::  :.:..     :  .   :  :..: .  .
      dektqmiyqqhqswlrpvlrsnrveycwcnsgraqckpiaekcseprcfnggtc_qqaly
      eekvchcqcaridwtaarcccklqv
            :::   .  :..::....
      fsdfv_cqcp__egfagkcceidt
```

L10 ANSWER 2 OF 29 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W19616 Protein DGENE
TITLE: New active mutants of radish antifungal protein 2 - used to generate fungus-resistant plants or as therapeutic or preservative agents
INVENTOR: Broekaert W F; De Samblanx G W; Rees S B
PATENT ASSIGNEE: (ZENE)ZENECA LTD
PATENT INFO: WO 9721814 A1 970619 39 pp
APPLICATION INFO: WO 96-GB3065 961212
PRIORITY INFO: GB 95-25474 951213
PAT. SEQ. LOC: Claim 1; Fig 1
DATA ENTRY DATE: 13 DEC 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-332785 [30]
DESCRIPTION: Radish antifungal protein 2 (Rs-AFP2)
KEYWORD: Rs-AFP2; radish antifungal protein 2; fungicide; salt tolerance; preservative; transgenic plant; crop protection
ORGANISM: Raphanus sativus
ABSTRACT:

This polypeptide comprises radish antifungal protein 2 (Rs-AFP2). Novel potent antifungal proteins (see W26371-90) based on Rs-AFP2 contain at least 1 mutation selected from a basic residue at positions 9 or 39, and a hydrophobic residue at positions 5 or 16. Proteins containing Gln5Met (see W26379), Gly16Met (W26380), Gly9Arg (W26376), Val39Arg (W26377) or Gly9Arg plus Val39Arg (W26378) mutations are specifically claimed. A cDNA clone encoding Rs-AFP2 preprotein can be modified by recombinant DNA methods to allow expression of mutant isoforms in yeast as mating factor alpha 1 fusion proteins. The Rs-AFP2 mutants have enhanced salt tolerant antifungal activity, especially when expressed in plant tissue where that may have curative as well as protective effects. They are useful for combating fungal diseases in agricultural, pharmaceutical or preservative applications

AMINO ACID COUNTS:3 A; 3 R; 5 N; 0 D; 0 B; 8 C; 3 Q; 1 E; 0 Z; 4 G; 2 H; 2 I; 2 L; 4 K; 0 M; 2 F; 3 P; 3 S; 1 T; 1 W; 2 Y; 2 V;
SEQUENCE LENGTH: 51
SEQUENCE

```
        1 qklcqrpsgt wsgvcgnnna cknqcirlek arhgscnyvf pahkcicyfp
       51 c
```

FEATURE TABLE:

Key	Location	Qualifier	
Misc_difference	5	note	"Gln at position 5 may be replaced by a hydrophobic amino acid, preferably Met""
Misc_difference	9	note	"Gly at position 9 may be replaced by a basic amino acid, preferably Arg""
Misc_difference	16	note	"Gly at position 16 may be replaced by a hydrophobic amino acid, preferably Met""
Misc_difference	39	note	"Val at position 39 may be

```
                |         |         |replaced by a basic amino
                |         |         |acid, preferably Arg""

ALIGN Smith-Waterman score:  88
   47 aa overlap starting at 2
   klascpegtavlscscgsacgswdireekvchcqcaridwtaarc_c
   :: . : ::    :. ..:: .  :: ::. : .:  . . : .: :
   klcqrpsgtwsgvcgnnnacknqcirlekarhgscnyv_fpahkcic
```

L10 ANSWER 3 OF 29 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W19281 Protein DGENE
TITLE: Antifungal peptide derived from radish antifungal protein 2 - and related DNA, useful for producing plants with increased fungal resistance and as therapeutic or preservative agent
INVENTOR: Borremans F A M; Broekaert W F; De Samblanx; Fant F; Meloen R H; Puijk W C; Rees S B; Schaaper W M M; Sijtsma L; Van Amerongen A; Van Gelder W M J
PATENT ASSIGNEE: (ZENE)ZENECA LTD
PATENT INFO: WO 9721815 A2 970619 65 pp
APPLICATION INFO: WO 96-GB3068 961212
PRIORITY INFO: GB 96-6552 960328
GB 95-25455 951213
PAT. SEQ. LOC: Disclosure; Figure 1
DATA ENTRY DATE: 21 JAN 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-332786 [30]
DESCRIPTION: Raphanus sativus antifungal protein 2 (Rs-AFP2)
KEYWORD: Antifungal protein; candida; fungal resistance; food additive; radish; crop protection; plant defensin; bacterial protection; preservative
ORGANISM: Raphanus sativus
ABSTRACT:

This protein sequence is the Rhapanus sativus (radish) mature antifungal protein (Rs-AFP2), which is homologous to proteins W19280- W19290. Shorter peptides, based on these proteins have been produced (see W19291-92, W19294-98, W19301-304, W19330-34 and W31765-834). Plants containing DNA sequences encoding these proteins have improved resistance to fungi. Compositions containing the peptides can be used to control fungi or bacteria in pharmaceutical (e.g. treatment of Candida infections) or preservative purposes (as food additives). In agriculture, the peptide may be used to improve disease resistance or disease tolerance of crops, either pre or post harvest. When applied to plants they may also have curative as well as protective actions. The peptides may also be used to protect plants by introducing them, or a microorganism capable of expressing the peptide into the soil

AMINO ACID COUNTS:3 A; 3 R; 5 N; 0 D; 0 B; 8 C; 3 Q; 1 E; 0 Z; 4 G; 2 H; 2 I; 2 L; 4 K; 0 M; 2 F; 3 P; 3 S; 1 T; 1 W; 2 Y; 2 V;
SEQUENCE LENGTH: 51
SEQUENCE

```
    1 qklcqrpsgt wsgvcgnnna cknqcirlek arhgscnyvf pahkcicyfp
   51 c
```

ALIGN Smith-Waterman score: 88
47 aa overlap starting at 2

```
klascpegtavlscscgsacgswdireekvchcqcaridwtaarc_c
:: . : ::    :. ..:: .  :: ::. : .:  . . : .: :
klcqrpsgtwsgvcgnnnacknqcirlekarhgscnyv_fpahkcic
```

L10 ANSWER 4 OF 29 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W26380 Protein DGENE
TITLE: New active mutants of radish antifungal protein 2 - used to generate fungus-resistant plants or as therapeutic or preservative agents
INVENTOR: Broekaert W F; De Samblanx G W; Rees S B
PATENT ASSIGNEE: (ZENE)ZENECA LTD
PATENT INFO: WO 9721814 A1 970619 39 pp
APPLICATION INFO: WO 96-GB3065 961212
PRIORITY INFO: GB 95-25474 951213
PAT. SEQ. LOC: Claim 2; Page 4
DATA ENTRY DATE: 13 DEC 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-332785 [30]
DESCRIPTION: Radish antifungal protein 2 mutant (G16M)
KEYWORD: Rs-AFP2; radish antifungal protein 2; fungicide; salt tolerance; preservative; transgenic plant; crop protection
ORGANISM: Chimeric Raphanus sativus; Chimeric synthetic
ABSTRACT:

This polypeptide comprises a specifically claimed Gly16Met mutant of radish antifungal protein 2 (Rs-AFP2) (see also W19616) that shows enhanced salt tolerant antifungal activity compared to the wild- type protein, especially when expressed in plants. A cDNA clone encoding Rs-AFP2 preprotein can be modified by recombinant DNA methods to allow expression of the mutant isoform in yeast as a mating factor alpha 1 fusion protein. Rs-AFP2 mutants (see also W26371-79 and W26381-90) are useful for combating fungal diseases in agricultural, pharmaceutical and preservative applications. When applied to plants, they may have curative as well as protective effects

AMINO ACID COUNTS:3 A; 3 R; 5 N; 0 D; 0 B; 8 C; 3 Q; 1 E; 0 Z; 3 G; 2 H; 2 I; 2 L; 4 K; 1 M; 2 F; 3 P; 3 S; 1 T; 1 W; 2 Y; 2 V;
SEQUENCE LENGTH: 51
SEQUENCE

```
       1 qklcqrpsgt wsgvcmnnna cknqcirlek arhgscnyvf pahkcicyfp
      51 c
```

ALIGN Smith-Waterman score: 86
47 aa overlap starting at 2

```
klascpegtavlscscgsacgswdireekvchcqcaridwtaarc_c
:: . : ::    :  ..:: .  :: ::. : .:  . . : .: :
klcqrpsgtwsgvcmnnnacknqcirlekarhgscnyv_fpahkcic
```

L10 ANSWER 5 OF 29 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W26379 Protein DGENE
TITLE: New active mutants of radish antifungal protein 2 - used to generate fungus-resistant plants or as therapeutic or preservative agents
INVENTOR: Broekaert W F; De Samblanx G W; Rees S B
PATENT ASSIGNEE: (ZENE)ZENECA LTD
PATENT INFO: WO 9721814 A1 970619 39 pp
APPLICATION INFO: WO 96-GB3065 961212
PRIORITY INFO: GB 95-25474 951213
PAT. SEQ. LOC: Claim 2; Page 4
DATA ENTRY DATE: 13 DEC 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-332785 [30]
DESCRIPTION: Radish antifungal protein 2 mutant (Q5M)
KEYWORD: Rs-AFP2; radish antifungal protein 2; fungicide; salt tolerance; preservative; transgenic plant; crop protection
ORGANISM: Chimeric Raphanus sativus; Chimeric synthetic
ABSTRACT:

This polypeptide comprises a specifically claimed Gln5Met mutant of radish antifungal protein 2 (Rs-AFP2) (see also W19616) that shows enhanced salt tolerant antifungal activity compared to the wild-type protein, especially when expressed in plants. A cDNA clone encoding Rs-AFP2 preprotein can be modified by recombinant DNA methods to allow expression of the mutant isoform in yeast as a mating factor alpha 1 fusion protein. Rs-AFP2 mutants (see also W26371-78 and W26380-90) are useful for combating fungal diseases in agricultural, pharmaceutical and preservative applications. When applied to plants, they may have curative as well as protective effects

AMINO ACID COUNTS:3 A; 3 R; 5 N; 0 D; 0 B; 8 C; 2 Q; 1 E; 0 Z; 4 G; 2 H; 2 I; 2 L; 4 K; 1 M; 2 F; 3 P; 3 S; 1 T; 1 W; 2 Y; 2 V;
SEQUENCE LENGTH: 51
SEQUENCE

```
        1 qklcmrpsgt wsgvcgnnna cknqcirlek arhgscnyvf pahkcicyfp
       51 c
```

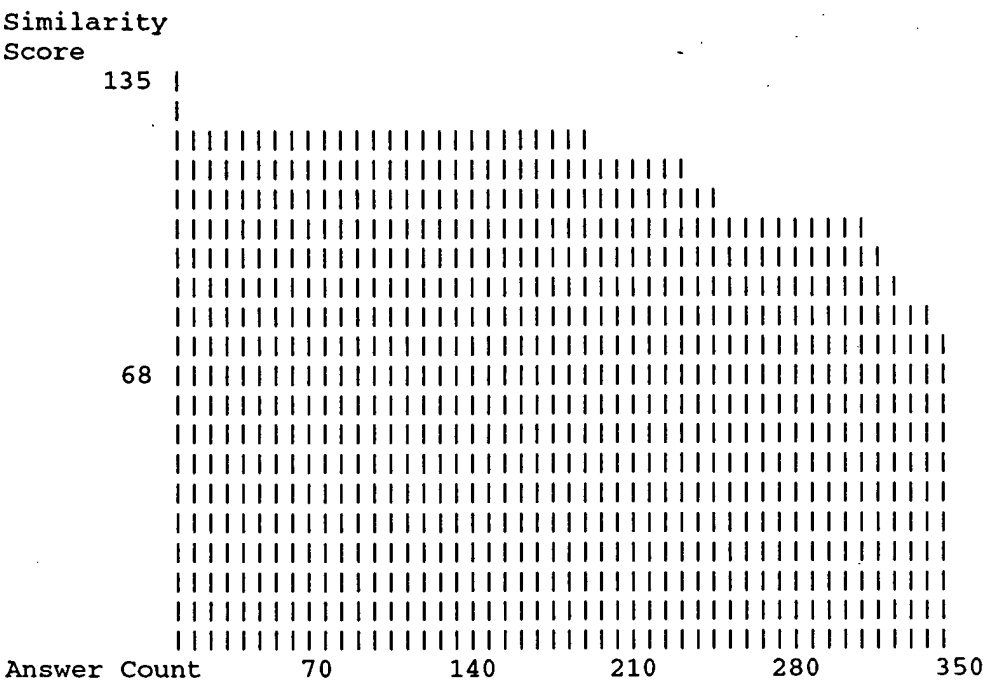
ALIGN Smith-Waterman score: 86
47 aa overlap starting at 2

```
klascpegtavlscscgsacgswdireekvchcqcaridwtaarc_c
::   : ::    :. ..:: .  :: ::. : .:  . . : .: :
klcmrpsgtwsgvcgnnnacknqcirlekarhgscnyv_fpahkcic
```

Rat C19

MKNLSFLLLFLFFLVLGLLGPSMSLCPMDEAISKKINQDFSSLLPAAMKNTVLHCWSVSSRGRLASCPEGTTVTSCSC
GSGCGSWDVREDTMCHCQCGSIDWTAARCCTLRVGS

333 ANSWERS FOUND ABOVE A THRESHOLD OF 68



L12 ANSWER 1 OF 333 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 87P-P70641 protein DGENE
TITLE: New modified tissue plasminogen activator - with new encoding DNA, new DNA expression vector, useful for treating vascular disorders, eg, pulmonary embolism arterial thrombosis
INVENTOR: Bang N U; Little S P; Schoner B E; Weigel B J
PATENT ASSIGNEE: (ELIL)ELI LILLY & CO
PATENT INFO: AU 8661804 A 870305 164 pp
APPLICATION INFO: AU 86-61804 860825
PRIORITY INFO: US 85-769298 850826
PAT. SEQ. LOC: Claim 9; page 122
DATA ENTRY DATE: 10 APR 1991 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 87-108842 [16]
CROSS REFERENCES: N-PSDB: 87N-N70990
DESCRIPTION: Modified tissue plasminogen activator
KEYWORD: Tissue plasminogen activator; kringle domain; embolism; thrombosis; stroke;
ORGANISM: Homo sapiens
ABSTRACT:

The modified t-PA has all/part of the kringle domains of native t-PA removed. The t-PA has functional properties superior to those of native t-PA. It retains fibrin binding properties and interacts more slowly and inefficiently with plasminogen activator inhibitor(s) compared to native t-PA. It is obtd. in large amts. from a prokaryotic host. Modified t-PA used for treating vascular disorders, eg deep vein thrombosis, pulmonary embolism, peripheral arterial thrombosis, disseminated intravascular coagulation, emboli from the heart or peripheral arteries, acute myocardial infarction, thrombotic strokes or fibrin deposits associated with invasive cancers. t-PA is used at a dosage of 250000 to 5000000 units at a loading dose or in a deep vein thrombosis-pulmonary embolism, or 250000-7500000 units over 30-90 mins. in acute myocardial infarction

AMINO ACID COUNTS: 18 A; 24 R; 9 N; 19 D; 0 B; 23 C; 22 Q; 19 E; 0 Z; 28 G; 12 H; 15 I; 30 L; 13 K; 5 M; 13 F; 19 P; 29 S; 16 T; 7 W; 13 Y; 20 V;
SEQUENCE LENGTH: 354
SEQUENCE

```
   1 mgsyqvicrd ektqmiyqqh qswlrpvlrs nrveycwcts graqchsvpv
  51 kscseprcfn ggtcqqalyf sdfvcqcpeg fagkcceist cglrqysqpq
 101 frikgglfad iashpwqaai fakhrrspge rflcggilis scwilsaahc
 151 fqerfpphhl tvilgrtyrv vpgeeeqkfe vekyivhkef dddtydndia
 201 llqlksdssr caqesslvrt vclppadlql pdwtecelsg ygkhealspf
 251 yserlkeahv rlypssrcts qhllnrtvtd nmlcagdtrs ggpqanlhda
 301 cqgdsggplv clndgrmtlv giiswglgcg qkdvpgvytk vtnyldwird
 351 nmrp
```

ALIGN Smith-Waterman score: 135
93 aa overlap starting at 1

```
lgpsmslcpmdeaiskkinqdfssllpaamkntvlhcwsvssrgrlascp_egttvtscs
.:  . .:  ::  .  .:  : : :.  .: : .:: .:.:..  : : .. .   :
mgsyqvic_rdektqmiyqqhqswlrpvlrsnrveycwctsgraqchsvpvkscseprcf
cgsgcgswdvredtmchcqcgsidwtaarcctl
 :. : .    : .:.:  :    :..:: .
nggtcqqalyfsdfvcqcpeg____fagkccei
```

L12 ANSWER 2 OF 333 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R74682 protein DGENE
TITLE: Genetically engineered tissue plasminogen activator - is modified at positions 44-50 and 296-302 and is non-glycosylated in K1 and K2 regions, has extended half-life and PAI-1 resistance
INVENTOR: Huang C; Huang P; Liu S
PATENT ASSIGNEE: (BIOE-N)BIOENGINEERING INST ACAD MILITARY
PATENT INFO: CN 1082111 A 940216 7 pp
APPLICATION INFO: CN 93-109234 930806
PRIORITY INFO: CN 93-109234 930806
PAT. SEQ. LOC: Claim 2; Fig 1 and Page 5
DATA ENTRY DATE: 04 JAN 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: Chinese
OTHER SOURCE: 95-162457 [22]
DESCRIPTION: t-PA mutein (N117Q, N184Q, delta 296-302, 44-50 replaced by ERHTSVQT)
KEYWORD: Tissue plasminogen activator; tPA; thrombolytic agent; mutein; deglycosylated kringle domain; PAI-1 resistance
ORGANISM: Synthetic
ABSTRACT:

The sequences given in R74678-R74689 are examples of preferred mutant versions of human tPA. In all the muteins, amino acids 296-302 of wild-type tPA (involved in interaction with PAI-1) have been deleted and the kringle domains have been deglycosylated by substn. of Asn 117 in K1 and Asn184 in K2 by Asp residues. Also, amino acids 44-50 of wild-type tPA are replaced by a motif which differs between different muteins. The modified tPA proteins have prolonged half-life, are resistant to PAI-1 and have affinity for fibrin; they are useful as thrombolytic agents

AMINO ACID COUNTS:32 A; 34 R; 20 N; 28 D; 0 B; 35 C; 29 Q; 27 E; 0 Z; 42 G; 15 H; 19 I; 39 L; 19 K; 5 M; 16 F; 27 P; 46 S; 27 T; 13 W; 24 Y; 24 V;
SEQUENCE LENGTH: 521
SEQUENCE

```
   1 syqvicrdek tqmiyqqhqs wlrpvlrsnr veycwcnsgr aqcerhtsvq
  51 tcseprcfng gtcqqalyfs dfvcqcpegf agkcceidtr atcyedqgis
 101 yrgtwstaes gaectnwqss alaqkpysgr rpdairlglg nhnycrnpdr
 151 dskpwcyvfk agkyssefcs tpacsegnsd cyfgqgsayr gthsltesga
 201 sclpwnsmil igkvytaqnp saqalglgkh nycrnpdgda kpwchvlknr
 251 rltweycdvp scstcglrqy sqpqfrikgg lfadiashpw qaaifaerfl
 301 cggilisscw ilsaahcfqe rfpphhltvi lgrtyrvvpg eeeqkfevek
 351 yivhkefddd tydndiallq lksdssrcaq essvvrtvcl ppadlqlpdw
 401 tecelsgygk healspfyse rlkeahvrly pssrctsqhl lnrtvtdnml
 451 cagdtrsggp qanlhdacqg dsggplvcln dgrmtlvgii swglgcgqkd
 501 vpgvytkvtn yldwirdnmr p
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	1..51	label	finger_domain
		note	"amino acids 44-50 of F domain have been replaced by the sequence ERHTSVQT"
Domain	52..88	label	E_domain

```
                |         |note     |"growth factor domain"
Domain          |89..177  |label    |Kringle_1
                |         |note     |"substn. of Asn117 (corresp.
                |         |         |to position 118 in this
                |         |         |mutein) by Asp destroys an
                |         |         |N-linked glycosylation site"
Domain          |178..276 |label    |Kringle_2
                |         |note     |"substn. of Asn184 (corresp.
                |         |         |to position 185 in this
                |         |         |mutein) by Asp destroys an
                |         |         |N-linked glycosylation site"
Domain          |277..521 |label    |P_domain
                |         |note     |"amino acids 296-302 of native
                |         |         |tPA have been deleted; these
                |         |         |residues are involved in
                |         |         |interaction with PAI-1"
Disulfide_bond  |6..36    |         |
Disulfide_bond  |34..43   |         |
Disulfide_bond  |52..63   |         |
Disulfide_bond  |57..74   |         |
Disulfide_bond  |76..85   |         |
Disulfide_bond  |93..174  |         |
Disulfide_bond  |114..156 |         |
Disulfide_bond  |145..169 |         |
Disulfide_bond  |181..262 |         |
Disulfide_bond  |202..244 |         |
Disulfide_bond  |233..257 |         |
Disulfide_bond  |265..389 |         |
Disulfide_bond  |301..317 |         |
Disulfide_bond  |309..378 |         |
Disulfide_bond  |403..478 |         |
Disulfide_bond  |435..451 |         |
   Disulfide_bond |468..496|         |

ALIGN Smith-Waterman score:  134
      87 aa overlap starting at 8
      deaiskkinqdfssllpaamkntvlhcwsvssrgrlascpegttvtscs___c__gsgcg
      ::  .   .:  : : :.  .: : .::  .: :: :.: . :.: .::   :  :. :
      dektqmiyqqhqswlrpvlrsnrveycw_cns_gr_aqcerhtsvqtcseprcfnggtcq
      swdvredtmchcqcgsidwtaarcctl
      .     : .:.:  :     :..:: .
      qalyfsdfvcqcpeg____fagkccei
```

L12 ANSWER 3 OF 333 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 98P-W54147 protein DGENE
TITLE: Mutant tissue plasminogen activator proteins - useful for treating vascular disorders, preventing tissue adhesion(s), etc
INVENTOR: Goeddel D V; Leung D W H; Rice G C
PATENT ASSIGNEE: (GETH)GENENTECH INC
PATENT INFO: US 5736135 A 980407 24 pp
APPLICATION INFO: US 95-389615 950213
PRIORITY INFO: US 91-728456 910711
US 93-8940 930126
US 94-221660 940401
US 95-389615 950213
PAT. SEQ. LOC: Claim 6; Page -
DATA ENTRY DATE: 20 JUL 1998 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 98-239153 [21]
DESCRIPTION: t-PA mutant (Y93C, T103A, N184S, G198D)
KEYWORD: Amino acid substitution; t-PA; vascular disorder; prevention; fibrin deposition; adhesion formation
ORGANISM: Synthetic
ABSTRACT:

Mutant tissue plasminogen activator proteins (W54147-W54158) are created by single or multiple amino acid substitutions. Compositions containing the t-PA variant are used for treating vascular disorders, for preventing fibrin deposition or for preventing adhesion formation or reformation. Note: This sequence is not given in the specification but was created from the wild type by the indexer

AMINO ACID COUNTS:33 A; 35 R; 21 N; 29 D; 0 B; 36 C; 26 Q; 26 E; 0 Z; 42 G; 16 H; 19 I; 39 L; 21 K; 5 M; 16 F; 29 P; 49 S; 24 T; 13 W; 34 Y; 14 V; 0 Others;
SEQUENCE LENGTH: 527
SEQUENCE

```
   1 syqvicrdek tqmiyqqhqs wlrpvlrsnr veycwcnsgr aqchsvpyks
  51 cseprcfngg tcqqalyfsd fvcqcpegfa gkcceidtra tccedqgisy
 101 rgawstaesg aectnwnssa laqkpysgrr pdairlglgn hnycrnpdrd
 151 skpwcyvfka gkyssefcst pacsegnsdc yfgsgsayrg thsltesdas
 201 clpwnsmili gkvytaqnps aqalglgkhn ycrnpdgdak pwchylknrr
 251 ltweycdyps cstcglrqys qpqfrikggl fadiashpwq aaifakhrrs
 301 pgerflcggi lisscwilsa ahcfqerfpp hhltyilgrt yrvvpgeeeq
 351 kfeyekyiyh kefdddtydn diallqlksd ssrcaqessv vrtvclppad
 401 lqlpdwtece lsgygkheal spfyserlke ahvrlypssr ctsqhllnrt
 451 ytdnmlcagd trsggpqanl hdacqgdsgg plyclndgrm tlygiiswgl
 501 gcgqkdypgy ytkvtnyldw irdnmrp
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	1..44	note	"Finger domain"
Domain	45..91	note	"Growth factor domain"
Domain	92..173	note	"Kringle-1 domain"
Domain	180..261	note	"Kringle-2 domain"
Domain	264..527	note	"Serine protease domain"
Misc_difference	93	note	"Y changed from wt to T in

```
                 |         |         |mutant"
Misc_difference  |103      |note     |"T changed from wt to A in
                 |         |         |mutant"
Misc_difference  |184      |note     |"N changed from wt to S in
                 |         |         |mutant"
Misc_difference  |198      |note     |"G changed from wt to D in
                 |         |         |mutant"                  -

ALIGN Smith-Waterman score:  129
      87 aa overlap starting at 8
      deaiskkinqdfssllpaamkntvlhcwsvssrgrlascp_egttvtscscgsgcgswdv
      ::  .   .:  : : :.  .: : .::  :.:..  : : .. .   :  :. : .
      dektqmiyqqhqswlrpvlrsnrveycwcnsgraqchsvpykscseprcfnggtcqqaly
      redtmchc______qcgsidwtaarcc
        : .:.:      .:  :: : : ::
      fsdfvcqcpegfagkcceid_tratcc
```

L12 ANSWER 4 OF 333 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R74688 protein DGENE
TITLE: Genetically engineered tissue plasminogen activator - is modified at positions 44-50 and 296-302 and is non-glycosylated in K1 and K2 regions, has extended half-life and PAI-1 resistance
INVENTOR: Huang C; Huang P; Liu S
PATENT ASSIGNEE: (BIOE-N)BIOENGINEERING INST ACAD MILITARY
PATENT INFO: CN 1082111 A 940216 7 pp
APPLICATION INFO: CN 93-109234 930806
PRIORITY INFO: CN 93-109234 930806
PAT. SEQ. LOC: Claim 2; Fig 1 and Page 5
DATA ENTRY DATE: 04 JAN 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: Chinese
OTHER SOURCE: 95-162457 [22]
DESCRIPTION: t-PA mutein (N117Q, N184Q, delta 296-302, 44-50 replaced by DNCRRPG)
KEYWORD: Tissue plasminogen activator; tPA; thrombolytic agent; mutein; deglycosylated kringle domain; PAI-1 resistance
ORGANISM: Synthetic
ABSTRACT:

The sequences given in R74678-R74689 are examples of preferred mutant versions of human tPA. In all the muteins, amino acids 296-302 of wild-type tPA (involved in interaction with PAI-1) have been deleted and the kringle domains have been deglycosylated by substn. of Asn 117 in K1 and Asn184 in K2 by Asp residues. Also, amino acids 44-50 of wild-type tPA are replaced by a motif which differs between different muteins. The modified tPA proteins have prolonged half-life, are resistant to PAI-1 and have affinity for fibrin; they are useful as thrombolytic agents

AMINO ACID COUNTS:32 A; 35 R; 21 N; 29 D; 0 B; 36 C; 28 Q; 26 E; 0 Z; 43 G; 14 H; 19 I; 39 L; 19 K; 5 M; 16 F; 28 P; 45 S; 25 T; 13 W; 24 Y; 23 V;
SEQUENCE LENGTH: 520
SEQUENCE

```
   1 syqvicrdek tqmiyqqhqs wlrpvlrsnr veycwcnsgr aqcdncrrpg
  51 cseprcfngg tcqqalyfsd fvcqcpegfa gkcceidtra tcyedqgisy
 101 rgtwstaesg aectnwqssa laqkpysgrr pdairlglgn hnycrnpdrd
 151 skpwcyvfka gkyssefcst pacsegnsdc yfgqgsayrg thsltesgas
 201 clpwnsmili gkvytaqnps aqalglgkhn ycrnpdgdak pwchvlknrr
 251 ltweycdvps cstcglrqys qpqfrikggl fadiashpwq aaifaerflc
 301 ggilisscwi lsaahcfqer fpphhltvil grtyrvvpge eeqkfeveky
 351 ivhkefdddt ydndiallql ksdssrcaqe ssvvrtvclp padlqlpdwt
 401 ecelsgygkh ealspfyser lkeahvrlyp ssrctsqhll nrtvtdnmlc
 451 agdtrsggpq anlhdacqgd sggplvclnd grmtlvgiis wglgcgqkdv
 501 pgvytkvtny ldwirdnmrp
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	1..50	label	finger_domain
		note	"amino acids 44-50 of F domain have been replaced by the sequence DNCRRPG"
Domain	51..87	label	E_domain

```
                |        |note     |"growth factor domain"
Domain          |88..176 |label    |Kringle_1
                |        |note     |"substn. of Asn117 by Asp
                |        |         |destroys an N-linked
                |        |         |glycosylation site"
Domain          |177..275|label    |Kringle_2
                |        |note     |"substn. of Asn184 by Asp
                |        |         |destroys an N-linked
                |        |         |glycosylation site"
Domain          |276..520|label    |P_domain
                |        |note     |"amino acids 296-302 of native
                |        |         |tPA have been deleted; these
                |        |         |residues are involved in
                |        |         |interaction with PAI-1"
Disulfide_bond |6..36    |         |
Disulfide_bond |34..43   |         |
Disulfide_bond |51..62   |         |
Disulfide_bond |56..73   |         |
Disulfide_bond |75..84   |         |
Disulfide_bond |92..173  |         |
Disulfide_bond |113..155 |         |
Disulfide_bond |144..168 |         |
Disulfide_bond |180..261 |         |
Disulfide_bond |201..243 |         |
Disulfide_bond |232..256 |         |
Disulfide_bond |264..388 |         |
Disulfide_bond |300..316 |         |
Disulfide_bond |308..377 |         |
Disulfide_bond |402..477 |         |
Disulfide_bond |434..450 |         |
   Disulfide_bond |467..495|         |

ALIGN Smith-Waterman score:  129
      83 aa overlap starting at 8
      deaiskkinqdfssllpaamkntvlhcwsvssrgrlascpe_gttvtscscgsgcgswdv
      ::  .   .:  : : :.  .: : .::  :.:..  .: . : .   :  :. : .
      dektqmiyqqhqswlrpvlrsnrveycwcnsgraqcdncrrpgcseprcfnggtcqqaly
      redtmchcqcgsidwtaarcctl
        : .:.:  :     :..:: .
      fsdfvcqcpeg____fagkccei
```

L12 ANSWER 5 OF 333 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R74678 protein DGENE
TITLE: Genetically engineered tissue plasminogen activator - is modified at positions 44-50 and 296-302 and is non-glycosylated in K1 and K2 regions, has extended half-life and PAI-1 resistance
INVENTOR: Huang C; Huang P; Liu S
PATENT ASSIGNEE: (BIOE-N)BIOENGINEERING INST ACAD MILITARY
PATENT INFO: CN 1082111 A 940216 7 pp
APPLICATION INFO: CN 93-109234 930806
PRIORITY INFO: CN 93-109234 930806
PAT. SEQ. LOC: Claim 2; Fig 1 and Page 5
DATA ENTRY DATE: 04 JAN 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: Chinese
OTHER SOURCE: 95-162457 [22]
DESCRIPTION: t-PA mutein (N117Q, N184Q, delta 296-302, 44-50 replaced by ESKPEAEE)
KEYWORD: Tissue plasminogen activator; tPA; thrombolytic agent; mutein; deglycosylated kringle domain; PAI-1 resistance
ORGANISM: Synthetic
ABSTRACT:

The sequences given in R74678-R74689 are examples of preferred mutant versions of human tPA. In all the muteins, amino acids 296-302 of wild-type tPA (involved in interaction with PAI-1) have been deleted and the kringle domains have been deglycosylated by substn. of Asn 117 in K1 and Asn184 in K2 by Asp residues. Also, amino acids 44-50 of wild-type tPA are replaced by a motif which differs between different muteins. The modified tPA proteins have prolonged half-life, are resistant to PAI-1 and have affinity for fibrin; they are useful as thrombolytic agents

AMINO ACID COUNTS:33 A; 33 R; 20 N; 28 D; 0 B; 35 C; 28 Q; 30 E; 0 Z; 42 G; 14 H; 19 I; 39 L; 20 K; 5 M; 16 F; 28 P; 46 S; 25 T; 13 W; 24 Y; 23 V;
SEQUENCE LENGTH: 521
SEQUENCE

```
      1 syqvicrdek tqmiyqqhqs wlrpvlrsnr veycwcnsgr aqceskpeae
     51 ecseprcfng gtcqqalyfs dfvcqcpegf agkcceidtr atcyedqgis
    101 yrgtwstaes gaectnwqss alaqkpysgr rpdairlglg nhnycrnpdr
    151 dskpwcyvfk agkyssefcs tpacsegnsd cyfgqgsayr gthsltesga
    201 sclpwnsmil igkvytaqnp saqalglgkh nycrnpdgda kpwchvlknr
    251 rltweycdvp scstcglrqy sqpqfrikgg lfadiashpw qaaifaerfl
    301 cggilisscw ilsaahcfqe rfpphhltvi lgrtyrvvpg eeeqkfevek
    351 yivhkefddd tydndiallq lksdssrcaq essvvrtvcl ppadlqlpdw
    401 tecelsgygk healspfyse rlkeahvrly pssrctsqhl lnrtvtdnml
    451 cagdtrsggp qanlhdacqg dsggplvcln dgrmtlvgii swglgcgqkd
    501 vpgvytkvtn yldwirdnmr p
```

FEATURE TABLE:

Key	Location	Qualifier	
Domain	1..51	label	finger_domain
		note	"amino acids 44-50 of F domain have been replaced by the sequence ESKPEAEE"
Domain	52..88	label	E_domain

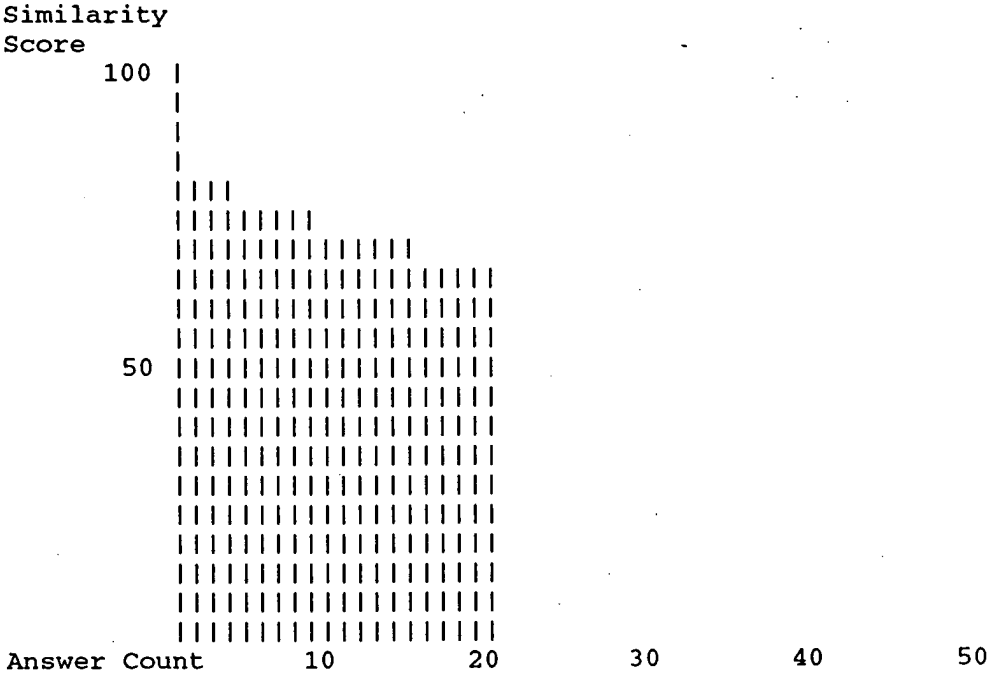
	|	|note	|"growth factor domain"
Domain	|89..177	|label	|Kringle_1
	|	|note	|"substn. of Asn117 (corresp.
	|	|	|to position 118 in this
	|	|	|mutein) by Asp destroys an
	|	|	|N-linked glycosylation site"
Domain	|178..276	|label	|Kringle_2
	|	|note	|"substn. of Asn184 (corresp.
	|	|	|to position 185 in this
	|	|	|mutein) by Asp destroys an
	|	|	|N-linked glycosylation site"
Domain	|277..521	|label	|P_domain
	|	|note	|"amino acids 296-302 of native
	|	|	|tPA have been deleted; these
	|	|	|residues are involved in
	|	|	|interaction with PAI-1"
Disulfide_bond	|6..36	|	|
Disulfide_bond	|34..43	|	|
Disulfide_bond	|52..63	|	|
Disulfide_bond	|57..74	|	|
Disulfide_bond	|76..85	|	|
Disulfide_bond	|93..174	|	|
Disulfide_bond	|114..156	|	|
Disulfide_bond	|145..169	|	|
Disulfide_bond	|181..262	|	|
Disulfide_bond	|202..244	|	|
Disulfide_bond	|233..257	|	|
Disulfide_bond	|265..389	|	|
Disulfide_bond	|301..317	|	|
Disulfide_bond	|309..378	|	|
Disulfide_bond	|403..478	|	|
Disulfide_bond	|435..451	|	|
Disulfide_bond	|468..496	|	|

```
ALIGN Smith-Waterman score:  129
      84 aa overlap starting at 8
      deaiskkinqdfssllpaamkntvlhcwsvssrgrlascpegttvt__scscgsgcgswd
      ::  .   .:  : : :.  .: : .::  :.:..  : ::.   .   :  :. : .
      dektqmiyqqhqswlrpvlrsnrveycwcnsgraqceskpeaeecseprcfnggtcqqal
      vredtmchcqcgsidwtaarcctl
         : .:.:  :     :..:: .
      yfsdfvcqcpeg____fagkccei
```

Human C23

MKALCLLLLPVLGLLVSSKTLCSMEEAINERIQEVAGSLIFRAISSIGLECQSVTSRGDLATCPRGFAVTGCTCGSAC
GSWDVRAETTCHCQCAGMDWTGARCCRVQP

20 ANSWERS FOUND ABOVE A THRESHOLD OF 66



L14 ANSWER 1 OF 20 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 97P-W24566 peptide DGENE
TITLE: Serine protease from Streptomyces griseus ATCC 55178 - with good stability in presence of urea or guanidine, useful in cleaning compositions, including laundry and dishwashing detergents
INVENTOR: Leigh S D
PATENT ASSIGNEE: (CLRX)CLOROX CO
PATENT INFO: US 5646028 A 970708 16 pp
APPLICATION INFO: US 91-718303 910618
PRIORITY INFO: US 91-718303 910618
US 92-973343 921106
US 94-292924 940818
US 95-544143 951017
PAT. SEQ. LOC: Claim 3; Column 25
DATA ENTRY DATE: 05 NOV 1997 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 97-362936 [33]
DESCRIPTION: Serine protease C-terminal sequence
KEYWORD: Serine protease; C-terminus; Streptomyces griseus; guanidine; pre-soak; cleaning composition; laundry detergent; additive composition; enzyme; dishwasher detergent; drain opener; urea; contact lens cleanser; proteinaceous stain
ORGANISM: Streptomyces griseus variety alkaliphilus No. 33
ABSTRACT:

This sequence represents the C-terminal sequence of the serine protease of the invention. The serine protease was isolated from Streptomyces griseus variety alkaliphilus No. 33 (ATCC 55178). The protease has an apparent molecular weight of 19 kD (by reducing sodium dodecylsulphate polyacrylamide gel electrophoresis), and improved stability against urea and guanidine. The serine protease is specific for the substrate represented by W24567, but also recognises the substrates shown in W26078-W26096. The protease is inhibited by phenylmethylsulphonyl fluoride. The serine protease is useful in liquid or granular cleaning compositions, specifically laundry detergents or additive compositions. It is also useful in automatic dishwasher detergents, pre-soaks, drain openers, contact lens cleansers etc. The protease has better activity against proteinaceous stains than known enzymes and unusually high stability in the presence of chaotropic agents

AMINO ACID COUNTS: 4 A; 4 R; 3 N; 1 D; 0 B; 4 C; 9 Q; 0 E; 0 Z; 18 G; 2 H; 6 I; 3 L; 0 K; 0 M; 2 F; 4 P; 15 S; 17 T; 1 W; 2 Y; 7 V;
SEQUENCE LENGTH: 102
SEQUENCE

```
        1 vtgstqatvg ssicrsgstt gwrcgtiqqh ntsvtypqgt itgvtrtsac
       51 aqpgdsggsf isgtqaqgvt sggsgncsig gttfhqpvnp ilsqygltlv
      101 rs
```

ALIGN Smith-Waterman score: 101
75 aa overlap starting at 6

```
qevagslifraissiglecqsvtsrgdlatcprgfavtgctcgsacgs_wdvraettchc
: ..:: : :. :. : .: .. ... .: :.: ..:: : ::::.. : .
qatvgssicrsgsttgwrcgtiqqhntsvtypqg_titgvtrtsacaqpgdsggsfisgt
qcagmdwtgarccrv
: :. :. : .
qaqgvtsggsgncsi
```

L14 ANSWER 2 OF 20 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 95P-R77256 Protein DGENE
TITLE: Pure, truncated fungal cellulase protein from Trichoderma - useful to reduce or eliminate dye, colourant or pigment back-staining or redeposition in stone-washing or bio-polishing
INVENTOR: Clarkson K A; Collier K D; Fowler T; Larenas E; Ward M
PATENT ASSIGNEE: (GEMV)GENENCOR INT INC
PATENT INFO: WO 9516782 A 950622 105 pp
APPLICATION INFO: WO 94-US14163 941219
PRIORITY INFO: US 93-169948 931217
PAT. SEQ. LOC: Claim 12; Page 38
DATA ENTRY DATE: 13 DEC 1995 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 95-231574 [30]
CROSS REFERENCES: N-PSDB: 95N-Q91276
DESCRIPTION: Truncated endoglucanase EGI catalytic core
KEYWORD: Cellulase; catalytic core; enzyme
ORGANISM: Trichoderma longibrachiatum
ABSTRACT:

A truncated fungal cellulase of Trichoderma comprising a EGI catalytic core with the sequence in R77256, which is encoded by Q91276, is claimed. The truncated cellulase is capable of endoglucanase activity. Genes for EGI and EGII have been isolated from T. longibrachiatum and the protein domain structure has been confirmed (Penttila, M. et al. 1986, Gene 45, 253-263; Van Arsdell, J.N. et al. 1987, Bio/Technology 5, 60-64; Saloheimo, M. et al., 1988, Gene 63, 11-21)

AMINO ACID COUNTS:2 A; 0 R; 2 N; 1 D; 0 B; 6 C; 4 Q; 0 E; 0 Z; 5 G; 0 H; 0 I; 1 L; 0 K; 0 M; 0 F; 2 P; 7 S; 2 T; 2 W; 3 Y; 2 V;
SEQUENCE LENGTH: 39
SEQUENCE

```
       1 qacssvwgqc ggqnwsgptc casgstcvys ndyysqclp
ALIGN Smith-Waterman score:  79
      13 aa overlap starting at 9
      qcagmdwtgarcc
      ::.:..:.:  ::
      qcgggqnwsgptcc
```

SEARCHED ON 26 OCT 1998
FILE LAST UPDATED: 18 OCT 1998 <19981018/UP>

L14 ANSWER 3 OF 20 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-R98208 Protein DGENE
TITLE: Cell-targetted retroviral vector particles - having envelope protein modified with targetting polypeptide
INVENTOR: Anderson W; Chiang Y L; Januszeski M; Mackrell A J; Zhao Y
PATENT ASSIGNEE: (GENE-N)GENETIC THERAPY INC
(UYSC-N) UNIV SOUTHERN CALIFORNIA
PATENT INFO: WO 9630504 A1 961003 73 pp
APPLICATION INFO: WO 96-US3908 960322
PRIORITY INFO: US 95-409648 950324
PAT. SEQ. LOC: Example 2; Page 36
DATA ENTRY DATE: 30 DEC 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-455352 [45]
DESCRIPTION: Nucleotide used in production of MSH/MoMuLV chimeric sequence
KEYWORD: Moloney murine leukaemia virus; gp70; 4070A retrovirus; retrovirus; 10A1 murine leukaemia virus; NZB-9-1 murine leukaemia virus; polytropic MX27 provirus; targetted drug delivery; gene therapy; single chain antibody; envelope protein; ss
ORGANISM: Synthetic
ABSTRACT:

Cell targetted retroviral vector particles can be used in gene therapy to deliver a heterologous gene to a target cell for expression of a heterologous polypeptide in that cell. The cell targetted retroviral vector particles comprise an envelope protein which is modified to contain a targetting polypeptide (a single chain antibody), or in the case of moloney murine leukaemia virus (MoMuLV), alpha melanotropin-stimulating hormone (MSH). Two oligonucleotides (R98207, R98208) were used to substitute sequences in MoMuLV for MSH sequences. This oligonucleotide was used to replace residues G80-P88 of MoMuLV envelope protein (See W04248)

AMINO ACID COUNTS:8 A; 0 R; 0 N; 0 D; 0 B; 17 C; 0 Q; 0 E; 0 Z; 8 G; 0 H; 0 I; 0 L; 0 K; 0 M; 0 F; 0 P; 0 S; 11 T; 0 W; 0 Y; 0 V;
SEQUENCE LENGTH: 44
SEQUENCE

```
        1 catttccgat ggtgcaagcc ggtattaacc tccctcaccc ctcg
   ALIGN Smith-Waterman score:  83
      43 aa overlap starting at 5
      tcprgfavtgctcgsacgswdvraettchcqcagmdwtgarcc
      ::  . . :::. :   :.  . :  :: : ::        ::
      tccgatggtgcaagccggtattaacctccctca_______cc
```

L14 ANSWER 4 OF 20 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 94P-R45359 Protein DGENE
TITLE: Transgenic plant contg. cDNA encoding wheat or barley lectin - has insecticidal properties in its leaves
INVENTOR: Raikhel N V
PATENT ASSIGNEE: (UNMS)UNIV MICHIGAN STATE
PATENT INFO: US 5276269 A 940104 26 pp
APPLICATION INFO: US 92-917665 920720
PRIORITY INFO: US 89-406318 890912
US 92-917665 920720
PAT. SEQ. LOC: Disclosure; Fig 6
DATA ENTRY DATE: 06 JUL 1994 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 94-016167 [02]
CROSS REFERENCES: N-PSDB: 94N-Q54433
DESCRIPTION: Wheat germ agglutinin isolectin WGA-D
KEYWORD: Transgenic plant; leaf; leaves; insecticidal; fungicidal; properties; tobacco; gramineae
ORGANISM: Triticum aestivum L
ABSTRACT:

The sequence is that of wheat germ agglutinin isolectin A, WGA-D, which may be expressed in transgenic plants to provide plants (pref. tobacco) having insecticidal and fungicidal properties in their leaves

AMINO ACID COUNTS:23 A; 5 R; 11 N; 5 D; 0 B; 32 C; 12 Q; 5 E; 0 Z; 42 G; 2 H; 3 I; 9 L; 8 K; 6 M; 6 F; 6 P; 16 S; 9 T; 3 W; 7 Y; 3 V;
SEQUENCE LENGTH: 213
SEQUENCE

```
  1 mrkmmstmal tlgaavflaf aaataqaqrc geqgsnmecp nnlccsqygy
 51 cgmggdycgk gcqngacwts krcgsqagga tcpnnhccsq yghcgfgaey
101 cgagcqggpc radikcgsqs ggklcpnnlc csqwgfcglg sefcgggcqs
151 gacstdkpcg kdaggrvctn nyccskwgsc gigpgycgag cqsggcdavf
201 agaitanstl lae
```

ALIGN Smith-Waterman score: 85
76 aa overlap starting at 13

```
gslifraissiglecqsvtsrgdlatcp_______rgf_avtgctcgsacgswdvraett
:. .: :...   . :    .:.   ::        :. .. :  ::..: .    .
gaavflafaaataqaqrcgeqgsnmecpnnlccsqygycgmggdycgkgcqngacwtskr
chcqcagmdwtgarcc
:  : .:    . .::
cgsqaggatcpnnhcc
```

L14 ANSWER 5 OF 20 DGENE COPYRIGHT 1998 DERWENT INFORMATION LTD
ACCESSION NUMBER: 96P-W02025 Protein DGENE
TITLE: Treatment of cellulose-contg. fabrics such as denim, e.g. stone:washing - using truncated cellulase enzyme to increase abrasion and give reduced redeposition of dye
INVENTOR: Clarkson K A; Collier K D; Fowler T; Larenas E; Ward M
PATENT ASSIGNEE: (GEMV)GENENCOR INT INC
PATENT INFO: WO 9623928 A1 960808 124 pp
APPLICATION INFO: WO 96-US977 960129
PRIORITY INFO: US 95-382452 950201
PAT. SEQ. LOC: Disclosure; Fig 2A-2E
DATA ENTRY DATE: 28 OCT 1996 (first entry)
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 96-371466 [37]
CROSS REFERENCES: N-PSDB: 96N-T32221
DESCRIPTION: Trichoderma cellobiohydrolase II
KEYWORD: Cellobiohydrolase II; CBHII; cellulase; cellulose; denim; stonewashing; dye redeposition; backstaining
ORGANISM: Trichoderma longibrachiatum
ABSTRACT:

The amino acid sequences for Trichoderma longibrachiatum cellobiohydrolase I (CBHI) (W02022), CBHII (W02025), endoglucanase I (EGI) (W02029), EGII (W02032) and EGIII (W02034) were deduced from the respective genomic DNA sequences (T32220-24). The CBHI, CBHII, EGI and EGII enzymes have catalytic core domains useful for reducing dye redeposition (backstaining) on cellulose-contg. fabrics such as denim, whilst maintaining or increasing abrasion during stonewashing. Truncated enzymes comprising these catalytic core domains can be obtd. by proteolysis of the complete enzyme or by inserting the appropriate DNA fragment into a vector, using this to transform a Trichoderma sp. host cell, and recovering the recombinant core domain

AMINO ACID COUNTS:60 A; 14 R; 30 N; 21 D; 0 B; 12 C; 21 Q; 10 E; 0 Z; 40 G; 4 H; 18 I; 38 L; 10 K; 5 M; 12 F; 32 P; 47 S; 38 T; 12 W; 20 Y; 27 V;
SEQUENCE LENGTH: 471
SEQUENCE

```
   1 mivgilttla tlatlaasvp leerqacssl wgqcggqnws gptccasgst
  51 cvysndyysq clpgaassss straasttsr vspttsrsss atpppgsttt
 101 rvppvgsgta tysgnpfvgv tpwanayyas evsslaipsl tgamataaaa
 151 vakvpsfmwl dtldktplme qtladirtan knggnyagqf vvidlpdrdc
 201 aalasngeys iadggvakyk nyidtirqiv veysdirtll viepdslanl
 251 vtnlgtpkca naqsayleci nyavtqlnlp nvamyldagh agwlgwpanq
 301 dpaaqlfanv yknasspral rglatnvany ngwnitspps ytqgnavyne
 351 klyihaigpl lanhgwsnaf fitdqgrsgk qptgqqqwgd wcnvigtgfg
 401 irpsantgds lldsfvwvkp ggecdgtsds saprfdshca lpdalqpapq
 451 agawfqayfv qlltnanpsf l
```

FEATURE TABLE:

Key	Location	Qualifier	
Peptide	1..24	label	Sig_peptide
Protein	25..471	label	Mat_protein
Domain	25..63	label	Cellulose_binding_domain

```
Region          |64..106 |label     |Linker_region
Domain          |107..471|label     |Catalytic_core_domain
                |        |note      |"catalytic core domain is the
                |        |          |preferred domain for use in
                |        |          |constructs of the invention"

ALIGN Smith-Waterman score:  82
      42 aa overlap starting at 4
      gfavtgctcgsacgswdvraettchc___qcagmdwtgarcc
      :. .:  : ..  .:  .. . .:     ::.:..:.:  ::
      gilttlatlatlaasvpleerqacsslwgqcgggqnwsgptcc
```